Edited by Timothy Shopen

Language Typology

and Syntactic Description

Volume I:
Clause Structure

SECOND EDITION

CAMBRIDGE

This page intentionally left blank

Language Typology and Syntactic Description

Second edition

Volume I: Clause Structure

This unique three-volume survey brings together a team of leading scholars to explore the syntactic and morphological structures of the world's languages Clearly organized and broad-ranging, it covers topics such as parts of speech, passives, complementation, relative clauses, adverbial clauses, inflectional morphology, tense, aspect, mood, and deixis. The contributors look at the major ways that these notions are realized, and provide informative sketches of them at work in a range of languages. Each volume is accessibly written and clearly explains each new concept introduced. Although the volumes can be read independently, together they provide an indispensable reference work for all linguists and field workers interested in cross-linguistic generalizations. Most of the chapters in the second edition are substantially revised or completely new – some on topics not covered by the first edition. Volume I covers parts-of-speech systems, word order, the noun phrase, clause types, speech act distinctions, the passive, and information packaging in the clause.

Timothy Shopen (1936–2005) was Senior Lecturer in Linguistics at the Australian National University. He had over forty years' experience of teaching and researching a variety of the world's languages, and also held posts at Indiana University and the Center for Applied Linguistics in Arlington, Virginia. In addition to *Language Typology*, he was editor of *Standards and Dialects in English* (1980), *Standards and Variables in English* (1981), *Languages and their Speakers* (1987), and *Languages and their Status* (1987).

Language Typology and Syntactic Description

Second edition
Volume I: Clause Structure

Edited by

Timothy Shopen[†]

Contents

Figures

Tables

Contributors

PAUL SCHACHTER, University of California, Los Angeles

TIMOTHY SHOPEN, Australian National University

MATTHEW S. DRYER, University at Buffalo

AVERY D. ANDREWS, Australian National University

EKKEHARD KÖNIG, Frei Universität, Berlin

PETER SIEMUND, Universität Hamburg

EDWARD L. KEENAN, University of California, Los Angeles

WILLIAM A. FOLEY, University of Sydney

Acknowledgements

Language typology studies what the languages of the world are like. When people ask 'What is linguistics?', from my point of view one of the best answers is 'the study of what the languages of the world are like'. I am honoured to have been joined by some excellent linguists in the achievement of this second edition of *Language Typology and Syntactic Description* for Cambridge University Press.

I am especially grateful to Matthew Dryer for coming in as co-editor when my health began to fail. Many thanks also to Lea Brown, for the invaluable help she gave Matthew in preparing the manuscript.

The Australian National University has always been generous in its support of my work. Except for the two and a half years I lived in Cairns, 2001 to 2003, it has been my base since I moved to Australia in 1975. I recognize the support I received from James Cook University during my time in Cairns.

I came up with the idea used to organize the first edition at a conference on field work questionnaires held at the Center for Applied Linguistics, Washington, DC. I said the best way to prepare for field work is to gain a good idea of what to look for. People thought this was right so I was asked to do the organizing. There have been surveys in the past but, I believe, none with this scope. The first edition has served as a reference manual and a textbook around the world and I have no doubt the second edition will as well. I have been pleased by the number of good linguists who have told me they have referred to our survey while doing field work valuable to us all.

Interest in the question of what the languages of the world are like is a longstanding one, but in the modern era Joseph Greenberg is an outstanding scholar who did important early work himself and was a model for others to do the same.

In an obituary for Joseph Greenberg by Steve Miller the distinction is made between taxonomists who are lumpers and splitters. Steve Miller says:

It is fitting that it was Darwin who first thought of the distinction between lumpers and splitters; the OED gives him the first citation of the words as applied to taxonomists. Lumpers gloss over or explain differences in pursuit of hidden unities; splitters do the opposite, stressing diversity.

Joseph Greenberg was a linguistic lumper and his dream of recreating the ur-language of humanity must stand as one of the greatest lumping dreams of all time. He dreamed of deep unity, and he spent an extremely long career pursuing evidence for it. He was still publishing highly technical evidence when he died, at age 85.

It is sad that he never published a manifesto, but he was a scientist and his inductive sensibility was not prone to making sweeping statements unsupported by minute attention to evidence. The nearest he came was in his conclusion to the controversial 1987 Language in the Americas, a book that grouped all languages in the western hemisphere into three families: 'The ultimate goal is a comprehensive classification of what is very likely a single language family. The implications of such a classification for the origin and history of our species would, of course, be very great.' Very great, as in, language was invented once and we might even have some ideas about what that language sounded like.

I was with Joseph Greenberg at Stanford University when he was doing his work, scouring through the part of the library that had grammars, making his counts: if you find construction x in a language you will always find, or you will be likely to find, construction y. This kind of commonality intrigued him. More from Steve Miller:

The splitters of linguistics have this problem: they're just not as interesting as the lumpers. The splitters' story is that the origins of language are irretrievable, so we should value every language for its expressive ability, but not for its place in the grand drama of linguistic diffusion. Greenberg, and the Nostraticists, and others who have tried to talk about language as a unity, dreamed something that may never be provable, but will continue to inspire us as a story that unites the human race as part of an ongoing story.

We give aid to both the lumpers and the splitters but, I believe, most of all to the lumpers. Languages differ from each other but only to a certain degree. Humankind is united in its use of language. This is an important message for us all as we go about our pursuits and combine with others to deal with the world.

Canberra, Australia
September 2004

TIMOTHY SHOPEN

Abbreviations and symbols

The following are abbreviations for grammatical terms used frequently in the glosses for examples. Other abbreviations are explained as they are presented.

A	subject of transitive clause
ABL	ablative
ABS	absolutive
ACC	accusative
ACT	actor
ADESS	adessive
ADJ	adjective
ADV	adverb
AFFIRM	affirmative
AG	agent
ALL	allative
ANIM	animate
ANT	anterior
ANTIPASS	antipassive
AOR	aorist
AP	actor pivot
APPLIC	applicative
ART	article
ASP	aspect
AUX	auxiliary
BEN	benefactive
BP	benefactive pivot
CAUS	causative
CF	counterfactual conjunction
CLSFR	classifier
COMP	complementizer
COMPAR	comparative
COMPLET	completive
CONCUR	concurrent

CONJ	conjunction
CONTIN	continuous
COP	copula
DAT	dative
DEBIT	debitive
DECL	declarative
DEF	definite
DEM	demonstrative
DEP	dependent
DET	determiner
DIR	directional
DIST	distal
DL	dual
DO	direct object
DP	directional pivot
DS	different subject
DU	dual
DUR	durative
EMPH	emphatic
ERG	ergative
ESS	essive
EXCL	exclusive
EXIST	existential
EXIST(NEG)	existential negator
EZ	ezafe
F, FEM	feminine
FUT	future
GEN	genitive
HABIT	habitual
HORT	hortative
IMPER	imperative
IMPERF	imperfect(ive)
IMPRS	impersonal
INAN	inanimate
INCEP	inceptive
INCL	inclusive
INCOMP	incompletive
INDEF	indefinite
INDIC	indicative
INFER	inferential
INFIN	infinitive
INSTR	instrumental

INT	interrogative
INTENS	intensifier
INTRANS	intransitive
INV	inverse
INVOL	involuntary
IO	indirect object
IRR	irrealis
LD	locative-directional
LINK	linker
LOC	locative
M MASC	masculine
MED	medial (intermediate between proximal and distal)
MID	middle
MOM	momentary
MONIT	monitory
MOT	motion
NC	noun class
NEAR	time close to now
NEG	negative
NEUT	neuter
NFN	non-finite
NOM	nominative
NOMIN	nominalization
NONLOCUT	nonlocutor
NONSPEC	nonspecific
NP	noun phrase
NUM	numeral
OBJ	object
OBLIQ	oblique
OBV	obviative
OM	object marker
OP	object pivot
OPT	optative
ORD	ordinal
P	object of transitive clause
PAR	partitive
PASS	passive
PAST	past tense
PC	paucal
PERF	perfect/perfective
PFX	prefix
PIV	pivot

PL	plural
PN	proper name / proper noun
POSS	possessive
POSSD	possessed
POTENT	potential
PP	prepositional phrase
PRED	predicative
PredP	predicate phrase
PREFL	possessive reflexive
PREP	preposition
PRES	present
PRET	preterite
PRO.ADJ	pro-adjective
PROG	progressive
PROHIB	prohibitive
PROL	prolative
PRT	particle
PTCL	particle
PTCPL	participle
PTV	primary transitive verb
PUNCT	punctual
PURP	purposive
Q	question marker
QUOT	quotative
REAL	realis
REC.PAST	recent past
RECIP	reciprocal
REFL	reflexive
REL	relative clause marker
REM	remote
REM.PAST	remote past
S	subject of intransitive clause
SEQ	sequential marker
SG	singular
SJNCT	subjunctive
S.O.	someone
SS	same subject
STAT	stative
SUBJ	subject
SUBORD SUFF	subordinative suffix
SUFF	suffix
SUPEREL	superelative

TNS	tense
TOP	topic
TR	transitive
TRANS	transitive
UNSPEC.OBJ	unspecified object
LONSPEC.SUBJ	unspecified subject
V	verb
VOL	volitional
VP	verb phrase
V.INTR	intransitive verb
V.DTR	ditransitive verb
V.TR	transitive verb
1	first person
2	second person
3	third person
1SG	first person singular (etc.)
3PL	third person plural (etc.)
ø	zero marking
-	affix boundary
=	clitic boundary
< >	infix
*	ungrammatical phrase or sentence
′	high tone
|	low tone
^	rise – fall tone or falling tone
??	only marginally grammatical
\	falling into nation contour [chapter 5]
/	rising into nation contour [chapter 5]

Unless otherwise indicated in a chapter, Roman numerals are used for noun classes.
'NC' with a subscript number 'x' means 'Noun class x'.

1 Parts-of-speech systems

Paul Schachter and Timothy Shopen[†]

0 Introduction

Parts of speech is the traditional term for the major classes of words that are grammatically distinguished in a language. While all languages make parts-of-speech distinctions, there are rather striking differences between languages with regard to both the kind and the number of such distinctions that they make. A field worker investigating an unfamiliar language may therefore find it useful to know what generalizations can be made about parts-of-speech systems. What, for example, can be said about the ways in which, and the limits within which, parts-of-speech inventories may differ from one another? Which parts-of-speech distinctions are universal and which language-specific? What are the ways in which languages that lack a particular part of speech express the semantic equivalent? And what relations are there between the parts-of-speech system of a language and the language's other typological characteristics? It is the aim of this chapter to provide some answers to such questions.

By way of orientation, the present section sets forth some general assumptions that underlie the presentation in the rest of the chapter. First, then, it is assumed here that the primary criteria for parts-of-speech classification are grammatical, not semantic. As has been amply demonstrated in the linguistic literature (see, for example, Fries (1952)), the familiar notional parts-of-speech definitions, such as 'a noun is the name of a person, place, or thing', fail to provide an adequate basis for parts-of-speech classification, since there are many cases in which their applicability or inapplicability is unclear. Grammatical criteria, on the other hand, are not open to this objection.

The grammatical properties of a word that are here taken to be relevant to its parts-of-speech classification include the word's distribution, its range of

syntactic functions, and the morphological or syntactic categories for which it is specifiable. Consider, in this connection, the three words of the sentence:

(1)   Boys like girls

The words *boys* and *like* can be shown to differ in their distributions (**Like boys girls* is ungrammatical), in their functional range (*boys* can function as a subject but *like* cannot) and in their categorizations *(boys* is categorized for number but not for tense, while *like* is categorized for both). Thus these two words are assigned to distinct parts-of-speech classes. On the other hand, the words *boys* and *girls*, having highly similar distributions (cf. *Girls like boys*), functional ranges, and categorizations, are assigned to the same parts-of-speech class. There are, to be sure, cases that are less clearcut than these – cases, for example, involving *partial* similarities of distribution, functional range, or categorization, which may require dividing a parts-of-speech class into subclasses. (For some further discussion, see section 1.) But, by and large, the grammatical properties in question constitute a serviceable basis for parts-of-speech classification.

While it is assumed here that the assignment of words to parts-of-speech classes is based on properties that are grammatical rather than semantic, and often language-particular rather than universal, it is also assumed that the *name* that is chosen for a particular parts-of-speech class in a language may appropriately reflect universal semantic considerations. Thus, although the familiar notional definition of nouns mentioned above does not always provide an adequate basis for deciding whether or not a given word is a noun, once the words of a language have been assigned to parts-of-speech classes on grammatical grounds and it is found that one of these classes includes the preponderance of words that are the names of persons, places, and things, then it is perfectly reasonable to call this class the class of nouns, and to compare the class so named with the similarly named classes of other languages. (On this point, see Lyons (1968:317–19).) Thus the words *boys* and *girls* are assigned to the same parts-of-speech class, and the word *like* to a different class, on language-particular grammatical grounds, but it is on universal semantic grounds that the class to which *boys* and *girls* are assigned is called the class of nouns, while that to which *like* is assigned is called the class of verbs.

An interesting recent proposal concerning universal semantic grounds for the identification of parts of speech is to be found in Wierzbicka (2000): namely, the use of universal exemplars, basic words that are presumably found in all languages, such as the equivalents of *person* and *thing* for nouns, *do* and *happen* for verbs. In any language, Wierzbicka suggests, the parts of speech that have been established on grammatical grounds that contain translations of these words can be said to be nouns and verbs respectively. She goes on to develop

this approach for a full range of parts of speech, arguing for the approach's superiority to other 'prototype approaches', which do not, she says, contain exemplars that are found in all languages. Notable exemplars are found in Dixon (1995), as well as Lyons (1977:vol. II), Croft (1984), Givón (1984a) and Hopper and Thompson (1984).

Another assumption reflected in this chapter is that all languages make a distinction between *open* and *closed* parts-of-speech classes. Following Robins (1964:230), we can describe open classes as those 'whose membership is in principle unlimited, varying from time to time and between one speaker and another' and closed classes as those that 'contain a fixed and usually small number of member words, which are [essentially] the same for all the speakers of the language, or the dialect'. Thus open classes are classes such as nouns and verbs, and closed classes are classes such as pronouns and conjunctions.

That all languages contain open classes is beyond doubt, despite occasional apocryphal reports to the contrary: i.e., reports of languages whose vocabularies consist of only a few hundred words. A more serious question can be raised about the universal status of closed classes. It is certainly true that closed classes play a rather minor role in some languages, and it has in fact sometimes been claimed that there are languages in which they play no role at all. The languages in question are invariably so-called *synthetic* languages: that is, languages that favour morphologically complex words. (Synthetic languages are commonly contrasted with analytic languages, in which words consisting of a single morpheme are the norm. If a scale were established, ranging from highly synthetic languages, such as Eskimo, to highly analytic ones, such as Vietnamese, modern English would be somewhat closer to the analytic than to the synthetic end of this scale.) The relation between a language's position on the synthetic–analytic scale and the role of closed classes in that language is discussed more fully in section 2. That section also considers, and rejects, the claim that there are known instances of languages with no closed classes at all.

The distinction between open and closed parts-of-speech classes provides the basic organizing principle of the remainder of this chapter, with open classes being dealt with in section 1 and closed classes in section 2.

1   Open classes

The open parts-of-speech classes that may occur in a language are the classes of *nouns*, *verbs*, *adjectives*, and *adverbs*. Typically, each of these classes may be divided into a number of subclasses on the basis of certain distinctive grammatical properties. For example, the class of nouns in English may be divided into such subclasses as common and proper (on the basis of whether or not the

nouns occur with articles like *the*: *the girl* vs **the Mary*), count and mass (on the basis of whether or not they occur in the plural: *chairs* vs **furnitures*), etc. And the class of English verbs may be divided into such subclasses as transitive and intransitive (on the basis of occurrence with objects: *enjoy it* vs **smile it*), active and stative (on the basis of occurrence in the progressive: *is studying* vs **is knowing*), etc. Such subclasses are not ordinarily identified as distinct parts of speech, since there are in fact properties common to the members of the different subclasses, and since the label *parts of speech* is, as noted earlier, traditionally reserved for 'major classes'. In any case, the discussion of open parts-of-speech classes in this chapter does not include a systematic account of the subclassification of these classes, but instead offers only a few observations concerning subclasses that are particularly widespread, or that seem particularly interesting from a typological viewpoint.

It must be acknowledged, however, that there is not always a clear basis for deciding whether two distinguishable open classes of words that occur in a language should be identified as different parts of speech or as subclasses of a single part of speech. The reason for this is that the open parts-of-speech classes must be distinguished from one another on the basis of a *cluster* of properties, none of which by itself can be claimed to be a necessary and sufficient condition for assignment to a particular class. And the fact is that languages vary considerably in the extent to which the properties associated with different open word classes form discrete clusters. Typically there is some overlap, some sharing of properties, as well as some differentiation. In English, for example, although nouns, verbs, and adjectives are clearly distinguished from one another in various ways, there are still certain properties that they share. Thus nouns and adjectives, as well as verbs, may be subclassified as active vs stative on the basis of occurrence in the progressive (compare *John is being a boor / boisterous* and **John is being my brother / tall*). And in certain other languages, as will become clear in the following sections, nouns and verbs, or nouns and adjectives, or verbs and adjectives, may have very much more in common than they do in English. What this means is that there may in some cases be considerable arbitrariness in the identification of two open word classes as distinct parts of speech rather than subclasses of a single part of speech. Thus some rather celebrated questions – for example, whether or not all languages make a distinction between nouns and verbs – may ultimately turn out to be more a matter of terminology than of substance (cf. section 1.2).

In the following presentation of the open parts-of-speech classes, nouns, verbs, adjectives, and adverbs are discussed in turn. In each case, the characteristic grammatical and notional properties of the class are enumerated, with relevant examples. Certain subclasses are also noted, and, where appropriate, there is a discussion of the question of the universality of a particular parts-of-speech distinction (see section 1.2), or of the ways in which languages that

lack a particular distinction express the semantic equivalent (see sections 1.3 and 1.4).

1.1 Nouns

The distinction between *nouns* and *verbs* is one of the few apparently universal parts-of-speech distinctions. While the universality of even this distinction has sometimes been questioned, it now seems that the alleged counter-examples have been based on incomplete data, and that there are no languages that cannot be said to show a noun–verb distinction when all relevant facts are taken into account. We shall look further into the matter of languages which allegedly fail to distinguish nouns and verbs at the end of section 1.2, after the characteristic properties of these two parts of speech have been described.

For convenience we can adapt the traditional definition of nouns, assigning the label *noun* to the class of words in which occur the names of most persons, places, and things. As was explained in the introductory section, this type of notional correlation is not the basis for determining membership in a class, but merely the basis for assigning a name to a class established on other grounds. It is therefore not a matter of concern if the class of nouns includes, as it typically does, words that are not the names of persons, places, or things, or if some such names are found in some other class.

It may be useful, however, to try to go beyond such traditional definitions to a deeper understanding of the semantics. Let us briefly consider in this connection some proposals made by Ronald W. Langacker (1987) and Anna Wierzbicka (1986). Langacker, working exclusively with English data, argues for certain universal semantic properties of nouns and verbs. Nouns, he proposes, do not foreground relations, but instead designate 'a region in some domain'. Verbs, on the other hand, do foreground relations. (For more on Langacker's views on verbs, see section 1.2, below.) Consider, for example, the difference in meaning between the following sentences:

(2)   The principal is speaking in the next room

(3)   The principal's speech is in the next room

The first sentence, using the verb *speak*, evokes an image of an audience and the principal communicating with them. The second sentence, using the noun *speech*, on the other hand, may simply serve to locate a physical entity; there is not necessarily any audience or any communication. Of course someone could read the speech (and the principal himself or herself could in fact read or recite it aloud to an audience), in which case communication would take place, but it could also be left unread with no communication. Thus the communication relation is not foregrounded.

Moreover, the speech in the context of the second sentence is likely to be understood as being in writing, which makes it easier to think of it as a 'region in some domain'. People can, of course, also say

(4)   The principal is giving a speech in the next room

in which case the noun necessarily represents something spoken and there is communication with an audience. In Langacker's scheme, however, this could be said to come from the force of the verbal expression with the verb *give*.

Langacker also makes an interesting suggestion about how the semantics of count and mass nouns can explain their syntactic differences. (Count nouns are nouns whose typical referents are countable and which may therefore be pluralized; mass, or noncount, nouns are nouns whose typical referents are not countable and which may therefore not normally be pluralized.) He suggests that count nouns can take plural inflection because their referents are 'bounded' in space. 'Bounded' means that, whether the count noun refers to a single entity (*dog*, *tree*) or to a set of entities (*crowd*, *herd*), the referent is conceived of as being defined in space. Mass nouns (*milk*, *sincerity*), on the other hand, refer to things that are conceived of as not 'bounded' but instead as having an indeterminate extent in space.

Like Langacker, Wierzbicka contrasts the semantic properties of nouns with those of another part of speech, but in her case this other part of speech is adjectives. Using examples from a variety of languages, she seeks to show how the semantics of nouns and adjectives can account for differences in how they are used. Nouns, she proposes, tend to refer to groupings of the permanent and/or conspicuous characteristics of entities. This is in contrast to adjectives, which tend to refer to a single temporary and/or less conspicuous characteristic. For example, to say *She is a cripple*, using the noun *cripple*, categorizes the person permanently, saying something about what kind of person she is. To say *She is sick*, with the adjective *sick*, on the other hand, says nothing about what kind of person she is, but instead refers to a single characteristic that the person has for the moment.

Because of these semantic differences, Wierzbicka argues, nouns are used for reference and categorization more easily than adjectives, while adjectives are used attributively more easily than nouns. Thus in a phrase such as *a sick woman*, the noun *woman* provides a broad categorization of the referent while the adjective *sick* serves to refine the categorization. By contrast, *a cripple woman*, with the two nouns providing a double categorization, is awkward.

An interesting example of the awkwardness of the attributive use of nouns is the following headline from the *Canberra Times* of 20 October 1999: *Diplomat murder accused granted bail*, which involves a recursive use of noun attributes. *Diplomat murder* means 'murder of a diplomat', *diplomat murder accused*

means 'person accused of murder of a diplomat', so the entire headline means 'bail has been granted to a person accused of the murder of a diplomat'.

The most common function for nouns is as arguments or heads of arguments – for example, as (heads of) subjects or objects, as in the case of the italicized words in:

(5)   The little *boy* was eating *candy*

Nouns may also function as predicates, however, either with an accompanying copula, such as English *be* (6) or Hausa *ne* (7), or without any copula, as in Tagalog (8) or Russian (9):

(6)   They are *teachers*

(7)   Su   *malamai* ne
       they teachers COP
       'They are teachers'

(8)   Mga *guro*   sila
       PL   teacher they
       'They are teachers'

(9)   Oni *učitelja*
       they teachers
       'They are teachers'

Typical categories for which nouns may be specified, either morphologically or syntactically, are case, number, class or gender, and definiteness. Case marking indicates grammatical functions (such as subject, direct object, and indirect object; cf. Andrews in chapter 3, and Dryer in chapter 4, of this volume for illuminating discussions of these functions), as in the following examples from Latin (10) (in which case is marked morphologically, by suffixation) and Japanese (11) (in which case is marked syntactically, by postpositions).

(10)   Femin-a   mal-um   puell-ae dedit
        woman-NOM apple-ACC girl-DAT gave
        'The woman gave an apple to the girl'

(11)   Onna   ga   shojo ni   ringo o   ataeta
        woman SUBJ girl   DAT apple OBJ gave
        'The woman gave an apple to the girl'

Number marking distinguishes singular from plural, and, more rarely, dual, as in English *house/houses*; Eskimo *iglu* 'house' / *iglut* 'houses' / *igluk* 'two houses'; or Tagalog *bahay* 'house' / *mga bahay* 'houses'. Class or gender marking partitions the set of nouns into subsets, each of which has its own distinctive marking and/or necessitates a distinctive marking on certain other words which

show agreement with nouns. Typically, the classification is in part semantically based and in part semantically arbitrary. Examples include the gender systems of Indo-European languages (e.g. German *der Mann* (the-masculine man) 'the man', *die Frau* (the-feminine woman) 'the woman', *das Mädchen* (the-neuter girl) 'the girl'), the class systems of Bantu languages (such as Swahili, in which most nouns that refer to human beings are in class I, which takes the prefix *m-*, e.g. *mtu* 'person', *mtoto* 'child', *mgeni* 'stranger', but in which some of the other classes have little semantic coherence), and the noun-classifier systems of such languages as Thai (cf. section 2.2, below). Some examples of definiteness distinctions are *a man* vs *the man*, Norwegian *en mann* 'a man' vs *mannen* 'the man', and Hebrew *ish* 'a man' vs *ha-ish* 'the man'.

In most languages some grammatical distinction is made between *common nouns*, which are used to refer to any member of a class of persons, etc. (e.g. *girl*, *city*, *novel*), and *proper nouns*, which are used to refer to specific persons, etc. (e.g. *Mary*, *Boston*, *Ivanhoe*). The precise character of the grammatical distinction, however, as well as its precise semantic correlates, may show considerable variation from language to language. For example, while common nouns in English differ from (most) proper nouns by occurring with articles, in Tagalog (which has no articles) common and proper nouns take different case markers and topic markers, as the following examples illustrate:

(12)   Malapit *sa*   babae   *ang* bata
         near   OBLIQ woman TOP child
         'The child is near the woman'

(13)   Malapit *kay*   Maria *si*   Juan
         near   OBLIQ Maria TOP Juan
         'Juan is near Maria'

Moreover, the Tagalog classes that are distinguished on this basis are not semantically coextensive with the English classes of proper and common nouns. The Tagalog nouns that take the markers of (13) are restricted to those that refer to specific *persons*; nouns that refer to specific *places*, etc., take the other set of markers, although their English equivalents are clearly proper, rather than common, nouns:

(14)   Malapit *sa*   Maynila *ang* Pasay City
         near   OBL Manila   TOP Pasay City
         'Pasay City is near Manila'

Apart from making a distinction between common and proper nouns, languages may make various other kinds of subclass distinctions within the set of nouns: for example, the distinction between count and mass nouns and the gender distinctions mentioned above. As was explained at the start of this

section, however, such subclass distinctions generally go beyond the scope of this chapter. (See Corbett in vol. III, chapter 4, on gender and noun class.)

1.2   *Verbs*

Verb is the name given to the parts-of-speech class in which occur most of the words that express actions, processes, and the like. As in the case of nouns, Langacker (1987) has proposed a deeper, more general account of the semantics, proposing, as noted above, that verbs, unlike nouns, foreground relations. In Langacker's scheme, however, the foregrounding of relations is not a unique property of verbs, since there are certain other parts of speech that can also foreground relations. What is distinctive about verbs, he suggests, is the foregrounding of *temporal* relations (relations that are anchored in time) or of relations concerned with process. Atemporal relations, on the other hand, are foregrounded by adpositions, adjectives, and adverbs (as well as by infinitives and participles, which Langacker does not classify as verbs).

The characteristic function of verbs is as predicates, as in:

(15)   The people *danced*
        The student *solved* the problem

In some languages, however, verbs can also occur as arguments as in the following example from Tagalog:

(16)   Pinanood ko   ang mga *sumasayaw*
        watch   I   TOP PL   were. dancing
        'I watched the ones who were dancing'
        cf. *Sumasayaw*   ang mga   tao
            were. dancing TOP PL   person
            'The people were dancing'

The use of a verb as an argument is to be distinguished from the probably more common use of a verbal *noun* as an argument, as in Akan:

(17)   Mehwɛɛ   *asaw*   no
        I. watched dancing the
        'I watched the dancing'

The verbal noun is a noun which is morphologically related to a verb, but which does not itself occur as a verbal predicate. For example, the verbal noun *asaw* of (17) is related to the verb *saw* 'dance' but could never itself be used as a predicate.

The categories for which verbs may be specified include tense, aspect, mood, voice, and polarity. (As in the case of nouns, the categorization may be manifested either morphologically or syntactically. Only morphological illustrations

will be given in this section, however. For some syntactic illustrations, see the presentation of auxiliaries in section 2.3. See also, for further information on tense, mood, and aspect, vol. III, chapter 5, by Timberlake, and, for a detailed treatment of mood, section 4 of that chapter.) Tense marking indicates time relative to the time of the utterance: for example Haya *akaija* 'he came (earlier than a few days ago)', *alaizile* 'he came (within the past few days)', *yaija* 'he came (earlier today)', *alaija* 'he will come (in the near future)', *aliija* 'he will come (in the distant future)'. Aspect marking indicates whether the action of the verb is regarded as complete or incomplete, durative or momentaneous, etc.: for example, Classical Greek *bebouleûsthai* 'to have already decided', *bouleúesthai* 'to be deciding', *bouleúseasthai* 'to decide' (unspecified for completeness or durativeness). Mood marking involves distinctions such as indicative (actual) vs subjunctive (possible) or declarative vs interrogative: for example, French *(qu')il viendra* '(that) he will come' vs *(qu')il vienne* '(that) he may come'; Menomini *pi·w* 'he is coming, he came' vs *pi·ʔ* 'is he coming?, did he come?'. Voice marking has to do with the role of the subject in the action expressed by the verb, the most common voice distinction being active vs passive, as in Latin *videt* 'he sees', *videtur* 'he is seen'. And polarity marking distinguishes affirmative from negative, as in Akan *tu* 'pulls', *ntu* 'doesn't pull'. (In addition to being marked for inherently verbal categorizations, verbs in some languages are marked to indicate certain categorizations (person, number, class) of their subjects and, less frequently, their objects: for example, Latin *video* 'I see', *videmus* 'we see'; Swahili *wa-ta-ni-uliza* (they-future-I-ask) 'they will ask me', *ni-ta-wa-uliza* (I-future-they-ask) 'I will ask them'.)

In all languages it is possible to subclassify verbs as transitive or intransitive on the basis of whether or not they occur with objects. In some languages the transitive–intransitive distinction entails certain other grammatical distinctions. For example, in Bambara the past tense is expressed by an auxiliary (*ye*) with transitive verbs but by a suffix (*-la*) with intransitive verbs:

(18)   U   *ye*   a   san
        they PAST   it buy
        'They bought it'

(19)   U   boli-*la*
        they walk-PAST
        'They walked'

Many languages also have a subclass of copulative verbs, like English *be*, that occur with predicate nominals or adjectives. In other languages, however, there is either no copula at all (as in Tagalog – cf. example (8)) or the copula is not a verb (as in Hausa – cf. example (7)). (For further discussion of non-verb copulas, see section 2.5.)

Another widespread subclassification of verbs involves a distinction between active verbs, which express actions and the like, and stative verbs, which express states and the like. With regard to these subclasses, once again Langacker (1987) makes an interesting suggestion about how the semantic properties of two subclasses can shed light on their formal properties. His account of the semantics of active and stative verbs (which he calls, respectively, perfective and imperfective) is, in fact, closely related to his account of the semantics of count and mass nouns (see section 1.1).

Specifically, Langacker proposes that active verbs (*walk*, *learn*) denote events conceived of as being bounded in time, much as count nouns denote entities conceived of as bounded in space. Stative verbs (*love*, *know*), by contrast, denote states of affairs conceived of as having an indeterminate extent in time, much as mass nouns denote entities conceived of as having an indeterminate extent in space. This leads him to an analysis of the situation in English where stative verbs are ordinarily used in the simple present tense for present time and active verbs in the present progressive (e.g., *John loves Mary* vs *John is walking to school*). When active verbs are used in the simple present tense, they ordinarily require a special interpretation because the bounded conception of the event cannot match with the extent of time of the speech act: habitual (*Ralph drinks two martinis for lunch*), imminent future (*The expedition leaves tomorrow at noon*), or historical present (*Then he walks up to me and says . . .*).

To turn now to the question of the universality of the noun–verb distinction, there are, as previously noted, languages with regard to which the legitimacy of such a distinction has been denied. Probably the best-known case is that of Nootka, which has often been cited in the linguistic literature as lacking a noun–verb distinction, on the basis of the analysis by Swadesh (1939). More recently, however, Jacobsen (1976) has re-examined the Nootka data, and has shown that, while the distinction between nouns and verbs in Nootka is less obvious than it is in many other languages, there is nonetheless a reasonably clear distinction to be made.

The following are the kind of examples that have been cited in support of the alleged lack of a noun–verb distinction in Nootka:

(20)   Mamu·k-ma   qu·ʔas-ʔi
         working-PRES(INDIC)   man-DEF
         'The man is working'

(21)   Qu·ʔas-ma   mamu·k-ʔi
         man-PRES(INDIC)   working-DEF
         'The working one is a man'

As these examples indicate, the notionally noun-like root meaning 'man', *qu·ʔas* and the notionally verb-like root meaning 'working', *mamu·k*, show, from the point of view of a language like English, rather surprising similarities of function

and categorizations. Thus *qu·ʔas* can function not only as an argument, as in (20), but also as a predicate, as in (21), without any accompanying copula. And *mamu·k* can function not only as a predicate but also as an argument (as in (20) and (21) respectively). Moreover, both the notionally noun-like and the notionally verb-like roots may be marked either for the typically nominal category 'definite' (by the suffix *-ʔi*) or the typically verbal category 'present' (by the suffix *-ma*).

What Jacobsen points out, however, is that the functional and categorizational ranges of roots like *qu·ʔas* and roots like *mamu·k*, although similar, are not identical. For example, while *qu·ʔas* and other notionally noun-like roots may function as arguments either with or without the suffix *-ʔi*, *mamu·k* and other notionally verb-like roots function as arguments only when suffixed. Compare (22) and (23):

(22)   Mamu·k-ma   qu·ʔas
        working-PRES(INDIC)   man
        'A man is working'

(23)   *Qu·ʔas-ma   mamu·k
        man-PRES(INDIC)   working

Moreover, some of the apparent similarities between nouns and verbs in Nootka turn out, on careful examination, to be of rather questionable significance. Thus there is evidence that Nootka tense morphemes, such as *-ma* in (20) and (21), are best analysed as clitics that attach to the clause-initial word, *whatever* category this word belongs to. (For fuller discussion, see Aikhenvald in vol. III, chapter 1.) It thus seems clear that Nootka does distinguish nouns and verbs, although this distinction is subtler than that found in English and many other languages.

Nootka is by no means alone, however, in making a fairly subtle distinction between nouns and verbs. Since the characteristic function of nouns is as arguments and that of verbs is as predicates, a functional distinction between nouns and verbs becomes difficult to establish to the extent that nouns occur as predicates and verbs as arguments without any distinctive marking (such as a copula accompanying the predicative nominal or some morpheme indicating nominalization of the verb). Consider in this connection the following examples from Tagalog:

(24)   Nagtatrabaho ang lalaki
        is. working   TOP man
        'The man is working'

(25)   Lalaki ang nagtatrabaho
        man   TOP is. working
        'The one who is working is a man'

As previously noted, and as further illustrated by (25), predicate nominals in Tagalog are not accompanied by a copula, and verbs occur freely as arguments. Thus, from a functional point of view, nouns and verbs appear to be at least as similar in Tagalog as they are in Nootka. (Tagalog does, however, make a more clearcut distinction in categorization: only verbs are inflectable for aspect.)

Nonetheless, while languages may differ considerably in the extent to which they make a grammatical distinction between nouns and verbs, it seems correct to say that all languages do in fact make some distinction between them. One might, however, wish to say that in some languages, such as Nootka and Tagalog, nouns and verbs have enough in common grammatically for there to be some question about whether to regard them as two subclasses of a single part of speech rather than two distinct parts of speech. Since this seems to be essentially a matter of terminology, it need not concern us further.

1.3   Adjectives

While all languages appear to distinguish two open classes, nouns and verbs, only certain languages make a further distinction between these and a third open class, the class of *adjectives*. The major question with which this section will be concerned is how adjectival meanings are expressed in languages that lack an open adjectival class. First, however, the properties of adjectives in those languages in which they do constitute a distinct open parts-of-speech class will be summarized.

The traditional notional definition of adjectives identifies them as the class of words denoting qualities or attributes. This definition has some well-known shortcomings (see, for example, the discussions in Jespersen (1924) and Lyons (1977)), but no obviously better notional definition has been proposed. As a result, even in notionally based grammars, adjectives have usually been defined at least in part in functional terms, as words which modify nouns. Among the words which modify nouns, a distinction is sometimes made between *limiting adjectives* and *descriptive adjectives*. However, the so-called limiting adjectives (words such as *some*, *this*, *other*) never constitute an open class, and will not here be treated as adjectives at all (see section 2.2 for a discussion of such words). The present discussion is thus confined to descriptive adjectives.

In addition to functioning as attributive modifiers of nouns (e.g. *tall* in *the tall woman*), adjectives may also function as predicates (as in *The woman is tall*). Like predicate nouns, predicate adjectives may or may not be accompanied by a copula. Thus English uses a copula while Ilocano does not:

(26)   Natayag daydyay babae
         tall   TOP   woman
         'The woman is tall'

A category for which adjectives are often specified is degree, which includes the traditional distinctions positive, comparative, and superlative – for example English *tall/taller/tallest*; Ilocano *natayag* 'tall' / *nataytayag* 'taller' / *katatayagan* 'tallest' – as well as various others: for example *very/too/so/rather tall*. In some languages, adjectives are also marked to indicate the categorizations of the nouns they modify or – when the adjectives are predicates – of the nouns that are their subjects. In Latin, for example, adjectives are marked for the case, gender, and number of nouns they modify (or are predicated of). Thus in

(27)   Feminae *procerae* homines *proceros* amant
         women   tall   men   tall   like
         'Tall women like tall men'

procerae is a nominative feminine plural form agreeing with *feminae* while *proceros* is an accusative masculine plural form agreeing with *homines*.

To turn now to the question of how the notional equivalent of adjectives is expressed in languages which lack an open adjective class, a distinction can be made between two groups of such languages. First, there are languages in which there is a class that can be called adjectives, but in which this class is closed rather than open, with anywhere from less than ten members (e.g. Igbo, which has eight) to fifty-odd (e.g. Swahili). And second, there are languages which lack a distinct adjective class altogether. Let us consider each of these groups in turn.

With regard to the first group, Dixon (1977b) has noted a rather striking cross-linguistic consistency in the range of meanings that the closed adjective class is used to express. Specifically, he finds that this class is likely to include words denoting dimensions (e.g. words meaning 'large' or 'small'), colour, age, and value (e.g. words meaning 'good' or 'bad'). On the other hand, it is less likely to include words denoting position ('high', 'low'), physical property ('hard', 'soft'), human propensity ('kind', 'cruel'), or speed. A paradigm case in support of Dixon's claim is offered by Igbo (see Welmers and Welmers (1969)), whose eight adjectives are neatly distributed among the four favoured semantic areas (see Table 1.1).

Table 1.1 *Igbo adjectives*

Dimension	Colour	Age	Value
ukwu 'large'	*ojii* 'black, dark'	*ọhụrụ* 'new'	*ọma* 'good'
nta 'small'	*ọca* 'white, light'	*ocye* 'old'	*ọjọọ* 'bad'

Dixon also suggests that there are some cross-linguistic tendencies, in languages with closed adjective classes, for certain specific types of 'adjectival' meanings to be expressed by verbs and other specific types by nouns. He suggests, for example, that physical properties are more often expressed by verbs than by nouns, while human propensities are more often expressed by nouns than by verbs. This seems, however, to be only a statistical tendency, and counterexamples are not hard to find. Thus Hausa seems to prefer nouns to verbs for expressing adjectival meanings in general, while Bemba seems to prefer verbs to nouns (see examples (28–33) below), although each of these languages sometimes uses the less favoured part of speech. In any event, it is clear that nouns and verbs between them must in general take up the slack left by a paucity of adjectives, and it is therefore of interest to see how each of these open classes is used to express adjectival meanings.

To begin with nouns, languages with closed adjective classes often use abstract nouns (equivalent in many cases to English nouns formed with -*ness*: *kindness*, *hardness*, etc.) in possessive constructions to express adjectival meanings. The following are some examples from Hausa, showing the syntactic parallelism between constructions with adjectival meanings and other constructions involving possessive modifiers (28), and possessive predicates (29):

(28)   mutum mai   alheri   /arziki   /hankali
         person having kindness /prosperity /intelligence
         'a kind/prosperous/intelligent person'
         cf. mutum mai   doki
             person having horse
             'a person having a horse'

(29)   Yana da   alheri   /arziki   /hankali
         he. is with kindness /prosperity /intelligence
         'He is kind/prosperous/intelligent'
         cf. Yana da   doki
             he. is with horse
             'He has a horse'

Some examples involving physical properties are:

(30)   itace mai   tauri   /laushi   /nauyi
         wood having hardness /softness /heaviness
         'hard/soft/heavy wood'

(31)   Yana da   tauri   /laushi   /nauyi
         it. is   with hardness /softness /heaviness
         'It is hard/soft/heavy'

The expression of adjectival meanings through *verbs* in languages with closed adjective classes typically involves relativization to express the equivalent of a modifying adjective. The following examples from Bemba are representative:

(32)   umuuntu   ùashipa   /ùakosa   /ùaceenjela
          person   who. is. brave /who. is. strong /who. is. wise
          'a brave/strong/wise person'
          cf. umuuntu   ùalemba
              person   who. is. writing
              'a person who is writing'

The equivalent of a predicate adjective, on the other hand, is expressed by a non-relativized verb:

(33)   Umuuntu áashipa /áakosa   /áaceenjela
          person   is. brave /is. strong /is. wise
          'The person is brave/strong/wise'
          cf. Umuuntu áalemba
              person   is. writing
              'The person is writing'

(As these examples indicate, relativized verbs in Bemba have low tone on the subject-concord prefix while non-relativized verbs have high tone on this prefix. In addition, with nouns in the human singular class – but not those in other classes – there is a segmental difference between the relative and non-relative subject prefixes.)

A further point to be noted about languages with closed adjective classes is that, in some of these languages, adjectives occur *only* as attributive modifiers, and do not occur as predicates at all. One such language is Hua, as is shown by the following examples (from Haiman 1978):

(34)   a. Bura fu nupa fu baie
             that pig black pig is
             'That pig is a black pig'

          b. *Bura fu nupa baie
              that pig black is

To return to the question of how adjectival meanings are expressed in languages that lack an open class of adjectives, let us now consider this question in relation to the second group of such languages: i.e., languages that have no distinct adjective class at all, either open or closed. Such languages can themselves be divided into two groups: languages in which adjectival meanings are expressed primarily by nouns (hereafter, *adjectival-noun languages*) and languages in which adjectival meanings are expressed primarily by verbs (hereafter, *adjectival-verb languages*).

In adjectival-noun languages, adjectival meanings seem in general to be expressed by nouns that designate an object (or objects) embodying a specified quality. The English equivalent of such nouns often takes the form adjective-plus-one(s), as the following example from the adjectival-noun language Quechua illustrates:

(35)   Rikaška: hatun-(kuna)-ta
        I saw   big-(PL)-ACC
        'I saw the big one(s)'

A comparison of (35) with (36) illustrates the grammatical similarity in Quechua between nouns with adjectival meanings (such as *hatun* 'big (one)') and other nouns (such as *alkalde* 'mayor').

(36)   Rikaška: alkalde-(kuna)-ta
        I saw   mayor-(PL)-ACC
        'I saw the mayor(s)'

As these examples show, nouns that express adjectival meanings can, like other nouns, be used as verbal objects, in which case they take the accusative suffix *-ta*, and can be pluralized by means of the suffix *-kuna*.

The following Quechua examples further illustrate the grammatical parallelism between nouns with adjectival meanings and other nouns in this adjectival-noun language:

(37)   Chay runa hatun (kaykan)
        that   man   big   (is)
        'That man is big'

(38)   Chay runa alkalde (kaykan)
        that   man   mayor   (is)
        'That man is mayor'

(39)   chay hatun   runa
        that   big   man
        'that big man'

(40)   chay alkalde runa
        that   mayor   man
        'that man who is mayor'

These examples show that nouns with adjectival meanings are not grammatically distinguished from other nouns either in their use as predicates or in their use as attributive modifiers. Thus, in (37) and (38), the predicates *hatun* and

alkalde both follow the subject and both optionally co-occur with the copulative verb *kaykan*, while in (39) and (40) the modifiers *hatun* and *alkalde* both immediately precede the noun they modify.

Adjectival-verb languages seem to be like languages with closed adjective classes in the way they use verbs to express adjectival meanings. As was noted above, in languages like Bemba, which have closed adjective classes but also use verbs extensively to express adjectival meanings, the usual verbal equivalent of a predicate adjective is a predicate verb in a non-relative construction (cf. example (33)), while the usual verbal equivalent of a modifying adjective is a verb in a relative construction (cf. example (32)). These same equivalents are found in adjectival-verb languages, as the following examples from one such language, Mandarin Chinese, illustrate:

(41)   Neige   nüaizi   piaoliang
         that   girl   beautiful
         'That girl is beautiful'

(42)   Neige   nüaizi   liaojie
         that   girl   understand
         'That girl understands'

(43)   piaoliang   de   nüaizi
         beautiful   REL   girl
         'a girl who is beautiful, a beautiful girl'

(44)   liaojie   de   nüaizi
         understand   REL   girl
         'a girl who understands, an understanding girl'

Examples (41) and (42) are predications, while examples (43) and (44) are modification constructions. As these examples show, verbs with adjectival meanings, such as *piaoliang* '(be) beautiful', and other verbs, such as *liaojie* 'understand', function in the same way in each of these construction types.

While there are some languages, such as Mandarin, which are clearly adjectival-verb languages, in that they appear to offer no consistent basis for distinguishing verbs with adjectival meanings from other verbs (or, at least, from other stative verbs such as 'understand' or 'know'), there are other languages whose classification as adjectival-verb languages is more problematic. These are languages in which the words that express adjectival meanings have most of the grammatical properties of (other) verbs – especially of stative verbs – but in which these adjectival words also have at least one distinctive property not shared by (other) verbs. One example of such a language is Mojave. In this language, adjectivals and stative verbs are indistinguishable when they are used

as predicates. Consider the following examples:

(45)   ʔi:pa-č   homi:-k   (iðu:m)
        man-SUBJ tall-PRES (AUX)
        'The man is tall'

(46)   ʔi:pa-č   su:paw-k   (iðu:m)
        man-SUBJ know-PRES (AUX)
        'The man knows'

As these examples illustrate, when used as predicates the adjectival stem *homi:* 'tall' and the stative verb stem *su:paw* 'know' take the same tense–aspect suffixes and optionally co-occur with the same auxiliary. (Non-stative verbs also take identical tense–aspect suffixes, but optionally co-occur with a different auxiliary.) Adjectivals are distinguishable from statives (and other verbs), however, when they are used as modifiers. When verbs are used as modifiers, they must appear in a relativized form, which in the relevant cases involves a prefixed k^w-, as in:

(47)   ʔi:pa k^w-su:paw-n^y-č   iva:k
        man   REL-know-DEM-SUBJ is. here
        'The man who knows is here'

(As the gloss indicates, the verb stem in this construction is followed by a demonstrative suffix and a case-marking suffix.) When adjectivals are used as attributive modifiers, on the other hand, the occurrence of the relativizing prefix is optional. Thus the following example is grammatical with or without the prefixed k^w-:

(48)   ʔi:pa (k^w-) homi:-n^y-č   iva:k
        man   (REL-)tall-DEM-SUBJ is. here
        'The tall man is here'

Compare the ungrammatical:

(49)   *ʔi:pa su:paw-n^y-č   iva:k
          man   know-DEM-SUBJ is. here

In the case of such a language, one would probably wish to analyse words with adjectival meanings as a distinguishable subclass of verbs rather than as a distinct part of speech, but this is perhaps an arbitrary choice.

1.4   Adverbs

Apart from nouns, verbs, and adjectives, there is one other open part-of-speech class that is attested in certain languages: the class of *adverbs*. The label *adverb*

is often applied to several different sets of words in a language, sets that do not necessarily have as much in common with one another, either notionally or grammatically, as, say, the subclasses of nouns or verbs that may occur in the language. For example, all of the italicized words of (50), which cover a considerable semantic and grammatical range, would ordinarily be identified as adverbs in a grammar of English:

(50)   *Unfortunately*, John walked *home extremely slowly yesterday*

A question may thus be raised as to whether there is sufficient similarity among the various types of 'adverbs' that may be recognized in a language to justify their being assigned to a single parts-of-speech class. We shall assume here that this question can, in general, be answered affirmatively, and that, for example, the italicized words of (50) can justifiably be assigned to a single parts-of-speech class, although they must obviously also be assigned to separate subclasses. (The subclass designations for these words would be 'sentence adverb' (*unfortunately*), 'directional adverb' (*home*), 'degree adverb' (*extremely*), 'manner adverb' (*slowly*), and 'time adverb' (*yesterday*). Some subclasses of adverbs may be closed rather than open, but since the class as a whole is open, it seems convenient to deal with the entire class in the present section.)

The usual functional definition of adverbs identifies them as modifiers of verbs, adjectives, or other adverbs (see, for example, Curme (1935)). In order to extend this definition so as to include sentence adverbs like *unfortunately* (which are in fact modifiers of entire sentences), and to allow for certain other possibilities (such as adverbs that modify entire verb phrases), we can say that adverbs function as modifiers of constituents other than nouns. The notional range of adverbs varies with the type of constituent modified. Sentence modifiers, for example, commonly express the speaker's attitude toward the event being spoken of; modifiers of verbs or verb phrases commonly express time, place, direction, manner, etc.; and modifiers of adjectives and adverbs commonly express degree.

Given the wide functional and notional range of adverbs, it is not surprising to find that there are no categorizations that are common to the entire class. In most cases, in fact, adverbs are not specified for any categories at all, although there are some exceptions. (Manner adverbs, for example, are sometimes specifiable for degree, as in *John worked hard/harder/hardest.*)

Some cross-linguistic observations may be made about the morphology of certain classes of adverbs. In many languages, manner adverbs are derivable from adjectives by means of fairly productive processes of derivational morphology. Thus in French many manner adverbs – as well as sentence and degree adverbs – are formed by adding the suffix *-ment* to the feminine singular form of an adjective: for example *lentement* 'slowly' (cf. *lente* 'slow (feminine singular)'), *malheureusement* 'unfortunately' (cf. *malheureuse* 'unfortunate (feminine singular)'), *activement* 'actively' (cf. *active* 'active (feminine

singular)'). And in Turkish many manner adverbs are formed by *reduplica-tion* of adjectives: for example *yavaş yavaş* 'slowly' (cf. *yavaş* 'slow'), *derin derin* 'deeply' (cf. *derin* 'deep'), '*acɨ acɨ* bitterly' (cf. *acɨ* 'bitter').

There is also a cross-linguistic tendency for manner adverbs – or a subset of manner adverbs – to have certain phonological properties that distinguish them from other words. For example, in Hausa many adverbs are high-tone monosyllables of the form obstruent-vowel-obstruent – for example, *kaf* 'completely', *kas* 'specklessly', *kat* 'with a snapping sound' – an otherwise rare pattern in this language. This phenomenon has received special attention in African linguistics, where the term *ideophone* has gained currency as the label for 'a word, often onomatopoeic, which describes a predicate, qualificative or adverb in respect to manner, colour, sound, smell, action, state or intensity' (Doke (1935:119)). But the phenomenon is, as noted by Courtenay (1976), by no means confined to African languages, and is attested, for example, in Australian languages as well. (Courtenay also notes that in some languages the peculiar phonological properties that distinguish ideophones from other words are not confined to adverbs. According to her analysis of Yoruba, for example, this language has ideophonic nouns, verbs, and adjectives, as well as adverbs. It seems, however, that while *all* adverbs in Yoruba are ideophonic, only relatively few nouns, verbs, and adjectives are.)

Before turning to consider how adverbial meanings are expressed in languages that lack a distinct open parts-of-speech class of adverbs, we should note that, even in languages with such a class, adverbial meanings are often expressible in other ways as well. In English, for example, phrases consisting of a preposition plus a noun or noun phrase can be used to express a wide range of adverbial meanings: time (*at dawn*), place (*in school*), direction (*to church*), manner (*with ease*), etc. And there are also expressions involving adjectives that paraphrase certain adverbs: for example *it is unfortunate that* (cf. *unfortunately*), *in a careless manner* (cf. *carelessly*). Not surprisingly, similar use is made of nouns and adjectives to express adverbial meanings in many languages in which there is no open adverbial class. (Some of the languages in question have a small, closed class of adverbs: others do not.)

In Arabic, for example, according to Bateson (1967), many adverbial meanings are expressed by nouns or adjectives (which Bateson considers a sub-class of nouns) in the accusative case. Relevant examples are *ɣadan* (next day: accusative) 'tomorrow' (cf. *ɣadu* 'next day'); *yoman* (day: accusative) 'daily' (cf. *yom* 'day'); *sariɛan* ('swift' accusative) 'swiftly' (cf. *sariɛ* 'swift'). In Tagalog, which lacks distinctive manner adverbs, the meaning equivalent of such adverbs is regularly expressed by adjectives preceded by the marker *nang*: *nang mabilis* (marker 'quick') 'quickly'; *nang malakas* (marker 'loud') 'loudly'; *nang bigla* (marker 'sudden') 'suddenly'; etc. There are also languages in which the meaning equivalent of a manner adverb is regularly expressed by an adjective without any special marking. One such language is Trique, in which,

according to Robert Longacre (personal communication), the class of adjectives simply does double duty, modifying verbs as well as nouns.

In some languages, the meaning equivalent of certain adverbs is expressed by *verbs*. This is particularly common in the case of comparative and superlative degree adverbs (e.g. English *more* and *most*), whose equivalent in a good many languages is expressed by a verb meaning 'surpass', as in the following Hausa examples:

(51)   a. Ya   fi   ni   hankali
            he.PERF   surpass   me   intelligence
            'He is more intelligent than I am'

      b. Ya   fi   su   duka   hankali
            he.PERF   surpass   them all   intelligence
            'He is the most intelligent of them all'

(There are also, of course, languages in which the comparative and superlative are expressed by affixes on adjectives, as in English *smarter*, *smartest*.) Some other examples of verbs expressing adverbial meanings are to be found in the following sentences from Akan:

(52)   a. ɔtaa   ba   ha
            he.pursue   come   here
            'He often comes here'

      b. Ohintaw   kɔ   hɔ
            he.hide   go   there
            'He goes there secretly'

The constructions in (52) are so-called *serial verb constructions* (see Schachter (1974) for some discussion).

Finally, it may be noted that in heavily synthetic languages, it is common for a wide range of adverbial meanings to be expressible by verbal affixes. Eskimo, for example, has a large set of suffixes with adverbial meanings that can occur between a verb root and an inflectional suffix. A few examples of such suffixes are: *-nirluk* 'badly', *-vluaq* 'properly', *-luinnaq* 'thoroughly', *-ɣumaaq* 'in the future', *-kasik* 'unfortunately', *-qquuq* 'probably'. And a similar situation obtains in Yana, where there are verbal suffixes such as *-ʔai* 'in the fire', *-xui* 'in(to) the water', *-sgin* 'early in the morning', *-ca* (*a*), 'at night', *-xkid* 'slowly', and *-ya*(*a*)*gal* 'quickly' (from Sapir and Swadesh (1960)).

2   Closed classes

Languages differ more from one another in the closed-class distinctions they recognize than in the open-class distinctions. This is true both of the *number*

and of the *type* of classes recognized. Thus there are languages which have been claimed (not quite correctly, as we shall see) to have no closed classes at all, while there are others that distinguish a dozen or more closed classes. And there may be no universally attested closed classes comparable with the universally attested open classes of nouns and verbs. (One closed class that is perhaps universal is the class of interjections – see section 2.5.) Nonetheless, it is apparently the case that, however diverse the closed-class systems of different languages may be, all languages do in fact have closed, as well as open, parts-of-speech classes.

Before we take a closer look at the kinds of closed parts-of-speech classes that occur, let us first consider the question of the correlation between the prominence of closed classes in a language and another typological feature: the position of the language on the analytic–synthetic scale. As was noted in the introduction to this chapter, languages may differ very greatly in the degree of morphological complexity they tolerate in words. Thus there are heavily analytic languages, in which there are few or no words that contain more than a single morpheme. And there are also heavily synthetic languages, in which polymorphemic words are the norm.

Not surprisingly, closed word classes tend to play a more prominent role in analytic languages than they do in synthetic languages. This is because much of the semantic and syntactic work done by the members of closed word classes in analytic languages is done instead by *affixes* in synthetic languages. We have already seen that, in some heavily synthetic languages, affixes may even do service for certain *open* word classes (cf. the Eskimo and Yana examples of affixal equivalents of adverbs cited at the end of the preceding section). The use of affixes in place of closed word classes is, however, a good deal more common – a claim that is substantiated in detail in the sections that follow. Therefore, by and large, the more use a language tends to make of morphologically complex words, the less use it will tend to make of closed word classes, and the fewer distinct types of closed classes it will tend to recognize.

It might therefore be expected that there would be some heavily synthetic languages that would make no use of closed classes at all. And in fact there have been claims to this effect with regard to at least one such language – witness the following quotation from Sapir (1921:119): 'In Yana the noun and verb are well distinct. But there are, strictly speaking, no other parts of speech.' However, subsequent investigation seems to have persuaded Sapir that, in addition to nouns and verbs, Yana does have a 'relational proclitic' (which is a kind of case marker, marking non-subjects) and a small set of articles (Sapir and Swadesh (1960)) – words which, in the terminology of this chapter, would be assigned to the closed class of *noun adjuncts* (see section 2.2 below). And it also seems likely that Yana has (or, rather, *had*, since the language is now extinct) some

interjections. Still, this is certainly a very meagre inventory of words belonging to closed classes, and it is safe to say that no analytic language could possibly manage with such an inventory.

In the discussion of closed word classes that follows, these classes are dealt with under the following headings: pronouns and other pro-forms (section 2.1), noun adjuncts (2.2), verb adjuncts (2.3), conjunctions (2.4), and other closed classes (2.5). These headings merely constitute a convenient framework for discussion, and are not claimed to have any theoretical status. In each case, the discussion of the closed classes in question will be accompanied by some discussion of the counterparts of these classes (if any) in languages in which the classes are not attested.

2.1   *Pronouns and other pro-forms*

This section surveys the various types of pro-forms that occur in languages and some of the ways in which languages that lack a particular pro-form type may express the semantic equivalent. The term *pro-form* is a cover term for several closed classes of words which, under certain circumstances, are used as substitutes for words belonging to open classes, or for larger constituents. By far the commonest type of pro-form is the *pronoun*, a word used as a substitute for a noun or noun phrase. Various subtypes of pronouns may be distinguished, among them *personal*, *reflexive*, *reciprocal*, *demonstrative*, *indefinite*, and *relative*. These subtypes are discussed in turn below. (There are also *interrogative pronouns*, but these are best considered together with other interrogative pro-forms – pro-adverbs, pro-verbs, etc. – and the discussion of them is thus deferred until later in the section.)

Personal pronouns are words used to refer to the speaker (e.g. *I*, *me*), the person spoken to (*you*), and other persons and things whose referents are presumed to be clear from the context (*he*, *him*, *she*, *her*, *it*, etc.). While personal pronouns in some languages occur in essentially the same sentence positions as other nominal expressions, it is rather common for them to show distributional peculiarities. This is true, for example, of direct-object personal pronouns in English, which must immediately follow the verb in some cases where other types of direct objects need not, as illustrated in (53).

(53)   Turn it on
        *Turn on it
        cf. $\begin{cases} \text{Turn the radio on} \\ \text{Turn on the radio} \end{cases}$

And it is more strikingly true of other languages in which personal pronouns are *clitics* whose distribution may be consistently distinct from that of non-clitic

nominals. (For further discussion of clitics, see section 2.5.) For example, object personal pronouns in French, both direct and indirect, normally precede the verb while non-pronominal objects normally follow it, as in (54).

(54)   a. Jean le leur   donnera
              Jean it to. them will. give
              'Jean will give it to them'

        b. Jean donnera le pain   aux   enfants
              Jean will give the bread to the children
              'Jean will give the bread to the children'

Similarly in Tagalog, personal-pronoun agents and topics normally follow the first constituent of the sentence, while other agents and topics normally follow the verb. For example:

(55)   Hindi ko   siya   nakita
        not   I(AG) him(TOP) saw
        'I didn't see him'
        cf. Hindi nakita ni Pedro si   Juan
             not   saw   AG Pedro TOP Juan
             'Pedro didn't see Juan'

It is quite common for the equivalent of personal pronouns, particularly of subject and object pronouns, to be expressed by affixes on the verb. The following examples of pronominal affixes are from Swahili and Quechua respectively:

(56)   Ni-li-wa-ona
        I-PAST-you-see
        'I saw you'

(57)   Maqa-ma-nki
        hit-me-you
        'You hit me'

Commonly such pronominal affixes may co-occur with non-affixal pronouns when a pronominal subject or object is being emphasized; compare (58) with (57):

(58)   Qam noqata maqamanki
        you me   hit. me. you
        'You hit *me'*

There are languages in which, while personal pronouns do occur, they are often avoided in favour of certain nouns which are considered to be more polite.

In Malay, for example (see Winstedt (1914)), a speaker under certain circumstances will use some self-deprecating noun (e.g. *hamba* 'slave') to refer to himself and some honorific noun (e.g. *tuan* 'master', *dato* 'grandfather', *nenek* 'grandmother') to refer to his addressee. (This situation led Robins (1964) to suggest that in Malay 'the nearest equivalents of the English pronouns are members of an open class' (p. 230). It seems, however, that while self-deprecating and honorific nouns are often used in place of pronouns, there are also unspecialized personal pronouns in Malay – *aku* 'I', *kamu* 'you', etc. – which may be used under appropriate circumstances.)

Finally it should be mentioned that in some languages the equivalent of a particular English personal pronoun may be expressed by the *absence* of any overt form in a particular context. In the Japanese sentences in (59) and (60), for example, there are no overt equivalents of 'he', 'her', and 'I'.

(59)   John wa   Mary o   sitte-imasu ga, amari yoku wa   sirima-sen
        John TOP Mary OBJ knows   but really well TOP knows-not
        'John knows Mary, but he doesn't know her very well'

(60)   Gohan o   tab-tai
        rice   OBJ eat-DESIDERATIVE
        'I want to eat rice'

Reflexive pronouns are pronouns which are interpreted as coreferential with another nominal, usually the subject, of the sentence or clause in which they occur. In English the reflexive pronouns are formed with *-self* or *-selves*:

(61)   John shaved *himself*
        John and Bill shaved *themselves*

It happens that in English *-self/-selves* pronouns are also used to indicate emphasis, as in:

(62)   John *himself* shaved Bill

In many languages, however, the reflexive and emphatic structures are formally unrelated, and the latter do not involve pronouns. Note, for example, the following Tagalog sentences:

(63)   Inahit   ni John ang *sarili niya*
        shaved AG John TOP self   his
        'John shaved himself'

(64)   Inahit   ni John *mismo* si   Bill
        shaved AG John EMPH TOP Bill
        'John himself shaved Bill'

(The emphatic *-self/-selves* forms and their non-pronominal counterparts in other languages should perhaps be considered a type of *discourse marker* – see section 2.2.)

Some languages which have distinct reflexive and non-reflexive third person pronouns do not make such a distinction for the other persons, but instead use the same first and second person pronouns both reflexively and non-reflexively. Note the following examples from French:

(65) a. Ils *me* voient
         they me see
         'They see me'

      b. Je *me* vois
         I   me see
         'I see myself'

(66) a. Ils *te* voient
         they you see
         'They see you'

      b. Tu *te* vois
         you you see
         'You see yourself'

(67) a. Ils *les* voient
         they them see
         'They see them'

      b. Ils *se* voient
         they REFL see
         'They see themselves'

There are also languages in which an invariable form is used for the reflexive, regardless of the person or number of the nominal with which it is coreferential. This is true, for instance, of Japanese, as the following examples illustrate:

(68) Taroo wa *zibun* o   mamotta
      Taroo TOP REFL OBJ defended
      'Taroo defended himself'

(69) Boku wa *zibun* o   mamotta
      I   TOP REFL OBJ defended
      'I defended myself'

(These examples, as well as some of those cited below, are from Faltz (1977).)

In a good many languages, reflexive forms are analysable as a head nominal modified by a pronominal possessive agreeing with the subject. Often the head nominal also occurs as a common noun meaning 'head' or 'body'. For example,

Fula reflexives, as in (70), are formed with *hoore* 'head', while Akan reflexives, as in (71), are formed with *ho* 'body'. There are also languages such as Malagasy (72), that use a common noun *without* a modifying possessive:

(70)   Mi gaañi   *hooreqam*
        I   wounded   my head
        'I wounded myself'

(71)   Mihuu   *me   ho*
        I saw   my   body
        'I saw myself'

(72)   Namono *tena* Rabe
        killed   body Rabe
        'Rabe killed himself'

   In languages that do not have reflexive pronouns or reflexively interpreted nouns or noun phrases, reflexive meanings may be expressed by verbal affixes, as in the following examples from Tswana (where the reflexive affix is *-i-*) and Lakhota (where it is *-ic'i-*):

(73)   Ke-tla-*i*-thêk-êla   selêpê
        I-FUT-REFL-buy-BEN axe
        'I shall buy an axe for myself'

(74)   Ophe*ic'i* thon
        'He bought it for himself'

        cf. Ophet^h on
             'He bought it'

   Reciprocal pronouns, like reflexive pronouns, are interpreted as coreferential with a co-occurring nominal, but are used to express mutual actions, conditions, etc. The English reciprocal pronouns are *each other* and *one another*, as in:

(75)   They helped *each other*
        They helped *one another*

Reciprocal and reflexive formations are often closely related. In Akan, for example, the reciprocal is formed by a kind of doubling of the reflexives:

(76)   Wohuu   wɔn ho   wɔn ho
        they. saw their body their body
        'They saw each other'

        cf. Wohuu   wɔn ho
            they. saw their body
            'They saw themselves'

There are, in fact, some languages in which there is regularly the possibility of ambiguity between reciprocal and reflexive meanings because the same forms can be used to express both. A case in point is French, as the following example illustrates:

(77)   Ils   *se*   flattent
        they   REFL/RECIP flatter
        'They flatter $\begin{Bmatrix} \text{themselves'} \\ \text{each other'} \end{Bmatrix}$

An unambiguously reciprocal meaning may, however, be conveyed by adding to (77) the phrase *l'un l'autre* 'one another', as in:

(78)   Ils   se   flattent l'un   l'autre
        they   REFL/RECIP flatter   the.one the.other
        'They flatter one another'

Languages that lack reciprocal pronouns, like those that lack reflexive pronouns, typically express equivalent meanings through the use of special affixes on the verb. In Ilocano, for example, reciprocal verbs contain the prefix *ag-* and the infix *-inn-* (which is inserted after the first consonant of the verb stem): for example *agsinnakit* 'hurt one another' (cf. *sakit* 'hurt'), *agtinnulong* 'help one another' (cf. *tulong* 'help').

Demonstrative pronouns are pronouns like English *this*, *that*, *these*, and *those* in:

(79)   *This* resembles *that*
        Do you prefer *these* or *those*?

Such pronouns are treated in depth by Comrie and Thompson in vol. III, chapter 6, and will only receive brief mention here. Demonstrative pronouns are widely attested. There are, however, languages in which demonstrative and third person personal pronouns are not distinguished. This is the case, for example, in Southern Paiute (see Sapir (1930)), where words consisting of a demonstrative morpheme followed by a third person morpheme do double duty as demonstrative and personal pronouns: for example *aŋa* (*a-* (that) + *-ŋa* (third person singular animate)) 'that one, he'; *iŋa* (*i-* (this) + *-ŋa*) 'this one, he'; *arï* (*a-* + *-rï* (third person singular inanimate)), 'that one, it', etc. (In addition to demonstrative pronouns, many languages have morphologically related demonstrative *articles*. For some discussion of these, see section 2.2.)

Indefinite pronouns are pronouns like English *someone*, *something*, *anyone*, *anything*. In many languages (including English) these forms are rather transparently analysable as consisting of two morphemes, one expressing the meaning of indefiniteness, the other the meaning 'person' or 'thing': for example Akan

obi (*o-* (human prefix) + *bi* (indefinite stem)) 'someone', *ebi* (*e-* (nonhuman prefix) + *bi)* 'something'; or French *quelq'un* (*quelqu'* (some) + *un* (one)) 'someone', *quelque chose* (*quelque* (some) + *chose* (thing)) 'something'.

Some languages have distinct indefinite-subject pronouns which are used to indicate an unspecified human subject. The English equivalent may have *they*, *you*, *one*, *people*, etc., according to the context. Some examples, from French (80), and Hausa (81), respectively, are:

(80)   a. *On*   dit   qu'il   pleut
          INDEF says that-it rains
          'They say that it's raining'

       b. *On*   ne   sait   jamais
          INDEF NEG knows never
          'You/One never can tell'

(81)   a. Kada   *a*   yi haka
          shouldn't INDEF do thus
          'You/One shouldn't do that'

       b. Yana so   *a*   zo
          he. is wanting INDEF come
          'He wants people to come'

Relative pronouns are pronouns like English *who* and *which* in:

(82)   The man *who* wrote that was a genius
       The book *which* he wrote was brilliant

Many languages do not have relative pronouns, but instead make use of *personal* pronouns in forming relative clauses, as in the following example from Akan:

(83)   Mihuu obi   a   ɔwɔ   aka   *no*
       I. saw someone REL snake has. bitten him
       'I saw someone whom a snake had bitten'
       cf. ɔwɔ   aka   no
           snake has. bitten him
           'A snake has bitten him'

Another common way of forming relative clauses involves *deletion* of the relativized nominal from the relative clause, as in the following example from Tagalog:

(84)   Sino ang bata-ng   pumunta sa   tindahan?
       who TOP child-LINK went   OBL store
       'Who is the child who went to the store?'

For still other relativization strategies, see Andrews in vol. II, chapter 4.

To turn now from pronouns to pro-forms of other types, the following types of pro-forms are discussed in turn below: *pro-sentences*, *pro-clauses*, *pro-verbs*, *pro-adjectives*, *pro-adverbs*, and *interrogative pro-forms*. This listing is not intended to be exhaustive. For example, it is argued in Schachter (1978) that the italicized words of sentences like (85) should be identified as *pro-predicates*:

(85)   Jack fell down, but Jill *didn't*
        Jill isn't crying, but Jack *is*

But this parts-of-speech type does not seem to be common enough to warrant discussion here.

Pro-sentences are words like English *yes* and *no*, which are used in answering questions, and which are understood as equivalent to affirmative and negative sentences respectively. (For example, in answer to *Is it raining?*, *Yes* is equivalent to *It's raining* and *No* to *It isn't raining.*) While most languages have such pro-sentences, they are not universal. In Mandarin Chinese, for example, the affirmative answer to a polar question is whatever verb occurred in the question, while the negative answer is *bu* 'not' optionally followed by this verb, as illustrated in (86).

(86)   Ni   qu ma?   Qu / Bu (qu)
        you go Q   go / not (go)
        'Are you going?'   'Yes'/ 'No'

There are also languages, however, which have a *larger* set of pro-sentences than English does. One rather common phenomenon is for *yes* to have two different equivalents, according to whether the question being answered is in the affirmative or in the negative. The following examples are from French:

(87)   Il   vient?   Oui
        he   comes   yes
        'Is he coming?'   'Yes'

(88)   Il   ne   vient   pas?   Si   (il   vient)
        he   NEG comes   not   yes (he   comes)
        'Isn't he coming?'   'Yes (he's coming)'

Another common phenomenon is the occurrence of a set of distinctive pro-sentences that are used in answer to existential questions (ones equivalent to English questions with *Is(n't) there?*, etc.). For example, in Tagalog the usual equivalents of *yes* and *no* are *oo* and *hindi* respectively:

(89)   Umuulan ba?   Oo   /Hindi
        is. raining Q
        'Is it raining?'   'Yes'   /'No'

In answer to existential questions, however, the existential pro-sentences *may-roon* and *wala* are used instead:

(90)   Mayroon ba-ng   pagkain?   Mayroon   /Wala
         EXIST   Q-LINK   food
         'Is there any food?'   'Yes'   /'No'

(As the question in (90) shows, *mayroon* is also used as an existential marker (cf. section 2.5) in non-pro-form sentences. The same is true of *wala*, which occurs as a negative existential marker.)

One common type of *pro-clause* is the question tag: a word with the force of a question which is appended to another clause. Some question tags are used to form *alternative* questions, others to form *confirmation* questions. Alternative questions are equivalent to certain English questions with *or* (e.g. *Is it raining or not?*), and quite commonly it is a word meaning 'or' that is used as an alternative-question tag, as in the following example from Hausa:

(91)   Ana   ruwa,   *ko*?
         one. is rain   or
         'Is it raining or not?'

Confirmation questions are questions in which the speaker is asking for confirmation of a statement to which the question tag is appended. An example, from Tagalog, is:

(92)   Umuulan,   *ano*?
         is. raining   CONFIRMATION TAG
         'It's raining, isn't it?'

In a good many languages, the equivalent of a question tag is expressed by a fixed idiomatic formula: for example French *n'est-ce pas* (literally 'is that not') or German *nicht wahr* (literally 'not true'). (In English the formula equivalent to a question tag is not fixed, but varies with the preceding statement: cf. *You haven't eaten, have you?*; *John left, didn't he?*; etc.)

Another type of pro-clause is the *so* or *not* of English sentences like

(93)   John says that it will rain, but I don't think *so*
         John says that it will rain, but I think *not*

(*So* and *not* in such cases are substitutes for *that* clauses: cf. *I don't think that it will rain*, *I think that it won't rain.*) In some languages, the same words that are

used as pro-sentences meaning 'yes' or 'no' can be used as pro-clauses. Thus in French, *oui* 'yes' and *non* 'no' occur in constructions like:

(94)   a. Je crois   que *oui*
            I   believe that yes
            'I believe so'

        b. Il dit   que *non*
           he says   that no
           'He says not'

Pro-verbs, *pro-adjectives*, and *pro-adverbs* are words which substitute for verbs (or verb phrases), adjectives (or adjective phrases), and adverbs (or adverb phrases) respectively. In Mandarin Chinese (see Chao (1968)), there are pro-verbs such as *lai* 'do it', *tzemme* 'do this', *nemme* 'do that'. An example, using the most common of these, *lai*, is:

(95)   Ni   buhui   xiu   zhe jigu,   rang wo *lai*
        you know.how.NEG repair this machine let   I   do. it
        'You don't know how to repair this machine; let me do it'

An example of a pro-adjective is the *le* of a French sentence such as (93).

(96)   Jean est grand, mais je ne   *le*   suis pas
        Jean is   tall   but I   NEG PRO.ADJ am   not
        'Jean is tall, but I'm not'

And an example of a pro-adverb is English *thus* or its (more commonly used) Akan equivalents *sɛɛ* 'this way' and *saa* 'that way, the same way':

(97)   Menoa no *sɛɛ*,   na   ɔno nso noa   no *saa*
        I. cook it this. way, and he too cooks it that. way
        'I cook it this way, and he cooks it the same way'

Interrogative pro-forms are words like English *who*, *what*, *where*, *when*, etc., as these are used at the beginning of questions. The set of interrogative pro-forms often cuts across other parts-of-speech classes. Thus in English there are interrogative pronouns (e.g. *who*, *what*), interrogative adverbs (e.g. *where*, *when*), and interrogative articles (e.g. *which* in *which book* – see section 2.2 for a general discussion of articles). And some other languages have interrogative pro-forms with no English counterparts. In Tagalog, for example, the interrog-ative root *ano* 'what' (which is also used as a confirmation-question tag) occurs

in various formations with adjective-forming and verb-forming affixes. Some examples of the resultant interrogative adjectives and verbs are:

(98)   a. *Napakaano* nila?
            very. what   they
            'What are they very much like?'
            cf. Napakataas nila
                 very. tall   they
                 'They are very tall'

        b. *Nagano*   ka?
            (PERF. ACTIVE)what   you
            'What did you do?'
            cf. Nagsalita   ka
                 (PERF. ACTIVE)speak you
                 'You spoke'

        c. *Naano*   ka?
            (PERF. INVOL)what   you
            'What happened to you?'
            cf. Natalisod   ka
                 (PERF. INVOL)trip   you
                 'You tripped'

It appears that all languages have interrogative pro-forms, but that the types of interrogative pro-forms that occur vary considerably from language to language, partly in conformity with the language's overall parts-of-speech system. Thus a language that lacks adverbs in general will naturally enough not have interrogative adverbs. For example, in Yana, which evidently has no adverbs, the equivalents of English *when*, *where*, etc., are all expressed by verb stems:

(99)   a. Beema'a-wara-nʒa-n?
            where-PERF-I-Q
            'Where am I?'

        b. Beeyauma-s aik nisaayau?
            when-FUT   his   going. away
            'When will he go away?'

2.2   *Noun adjuncts*

This section discusses several classes of words that typically form phrasal constituents with nouns. In most cases, these words, here labeled *noun adjuncts*, have clear semantic import, conveying some information about the referent of the phrasal constituent that is not expressed by the noun itself: for example the role of the referent in the action expressed by a co-occurring verb, or whether

the referent is singular or plural. In some cases, however, the noun adjuncts appear to be semantically empty and merely to be required under certain circumstances by the syntax of the language. (This seems to be true, in particular, of certain *classifiers* – see below.) Four general classes of noun adjuncts will be distinguished here: *role markers*, *quantifiers*, *classifiers*, and *articles*. These classes are discussed in turn below.

Role markers include *case markers*, *discourse markers*, and (other) *adpositions* (i.e., *prepositions* or *postpositions*). Case markers are words that indicate the syntactic and/or semantic role (e.g. subject and/or agent) of the noun phrase to which they belong. Discourse markers are words that indicate the discourse role (e.g. topic) of the associated noun phrase. If the role marker precedes the noun, as in the following example from Tagalog, it may be called a preposition:

(100)   Ipinansulat *ni* John *ng* liham *kay* Mary *ang* makinilya
         wrote. with AG John OBJ letter IO Mary TOP typewriter
         'John wrote Mary a letter on the typewriter'

And if the role marker follows the noun, as in the following example from Japanese, it may be called a postposition:

(101)   Type   *de*   *wa* John *ga* Mary ni tegami *o*   kaita
         typewriter INSTR TOP John SUBJ Mary IO letter   OBJ wrote
         'John wrote Mary a letter on the typewriter'

There are certain adpositions that are clearly not discourse markers, but that are not ordinarily identified as case markers: for example the postposition *de* in (101), and the words indicating various locative relations in the following examples from English and Akan respectively:

(102)   It's *on/under/beside* the table

(103)   ɛwɔ pon   no *so/ase/nkyɛn*
         it. is table the on/under/beside
         'It's on/under/beside the table'

The distinction between case-marking and other adpositions seems to be a somewhat arbitrary one, however, based in part on the traditional identification of only certain types of grammatical or semantic roles with the label 'case'. (It is also true, however, that in some languages locative and other adpositions must be distinguished from case-marking *affixes* which occur in the same phrase. For example, in Latin 'on the table' is *super mensam*, which consists of the preposition *super* 'on' and a case-marked form, the accusative, of the noun *mensa* 'table'. In such cases the adposition shows the relation of a noun (phrase) to some larger syntactic unit while the affix shows its relation to the adposition.)

A certain correlation is known to exist between the general word order type of a language and the occurrence in the language of prepositions as opposed to postpositions. In particular, Greenberg (1963:62) has claimed that verb-initial languages are always prepositional while verb-final languages are almost always postpositional (cf. (100) and (101) above). In verb-medial languages the situation is less clearcut. While the majority of such languages are prepositional (cf. (102)), there are also a good many that are postpositional instead (cf. (103)).

In languages that do not use role markers to indicate the grammatical, semantic, or discourse roles of nouns, or that use markers for some such roles but not others, the roles in question may be indicated by word order or by affixation. English, for example, uses word order to distinguish subjects from objects (compare *The boy loves the girl* and *The girl loves the boy*) while languages like Latin and Warlpiri use case-marking. Compare (104) and (105) for Warlpiri with free word order and a case-marking clitic for Ergative (*-ngku*), transitive subject. The case-marking clitic can attach itself to every word in a noun phrase or just the last word of the noun phrase:

(104)   a. Ngarrka-ngku karnta   nyangu
           man-ERG   woman saw
           'The man saw the woman'

        b. Nyangu ngarrka-ngku karnta
           saw   man-ERG   woman
           'The man saw the woman'

(105)   a. Ngarrka nyangu karnta-ngku
           man   saw   woman-ERG
           'The woman saw the man'

        b. Ngarrka karnta-ngku nyangu
           man   woman-ERG saw
           'The woman saw the man'

There are also languages in which the affixes that indicate the role of a noun or noun phrase may appear on a *verb* rather than on the noun or noun phrase itself. In Swahili, for example, the affix *-i-* on a verb expresses the equivalent of the benefactive preposition *for*, as in (106).

(106)   Ni-li-m-p-*i*-a   chakula mwanamke
        I-PAST-her-give-for-ø food   woman
        'I gave food for the woman'
        cf. Ni-li-m-p-a   chakula mwanamke
            I-PAST-her-give-ø food   woman
            'I gave the woman food'

(See also the Tagalog example (100), above, in which the affix *ipinan-* on the verb *ipinansulat* indicates that the topic noun phrase *ang makinilya* is to be interpreted as playing the role of an instrument.)

Finally, it may be noted that certain discourse roles are sometimes indicated by the use of special syntactic constructions or by intonation. Thus the English equivalent of the Akan focus marker *na* of (107) is the so-called cleft-sentence construction, while the English equivalent of the Akan contrast marker *de* of (108) is intonational:

(107)   Kwame *na*   ɔbɛyɛ   adwuma no
         Kwame FOCUS he. will. do work   the
         'It's Kwame who will do the work'

(108)   Kwame *de*,   ɔbɛkɔ   na Kofi de,   ɔbɛtena   ha
         Kwame CONTRAST he. will. go and Kofi CONTRAST he. will. stay here
         '*Kwame* will go but *Kofi* will stay here'

The next group of noun adjuncts to be considered, the *quantifiers*, consists of modifiers of nouns that indicate quantity or scope: for example numerals, and words meaning 'many', 'much', 'few', 'all', 'some', 'each', etc. In some languages a quantifier is required if plurality is to be explicitly indicated. Tagalog, for example, uses the quantifier *mga* in this way:

(109)   Nasaan ang *mga* pinggan?
         where   TOP PL   dish
         'Where are the dishes?'

And Vietnamese appears to have some fifty different pluralizers (which, however, also carry some more explicit quantifier meaning: 'all', 'all (vague)', 'many', 'a few', etc. – see Binh (1971:113–14)). In such languages the explicit indication of plurality is generally optional. Thus if *mga* is deleted from (109), the resultant sentence can still be interpreted as meaning 'Where are the dishes?', although it could also mean 'Where is the dish?'

There are a number of languages in which quantifiers, or at least certain quantifiers, vary in form according to the semantic properties of the nouns they modify. Thus the Akuapem dialect of Akan has two distinct forms for certain numerals, according to whether the noun modified is human or nonhuman: for example *nnipa baanu* (people two) 'two people' vs *mmoa abien* (animals two) 'two animals'. And in Japanese there are semantically conditioned variants such as *sannin* 'three (of humans)', *sanba* 'three (of birds)', *sanbon* 'three (of cylindrical objects)', *sanmai* 'three (of thin flat objects)', etc. (The Japanese examples are bimorphemic, each consisting of the quantifier morpheme *san-* 'three' plus a classifier morpheme: see the discussion of classifiers below.)

The range of meanings expressed by a distinct parts-of-speech class of quantifiers varies considerably from language to language, and languages that have such a class may nonetheless have other means for expressing particular quantity or scope meanings. One such means involves nouns of quantity or scope in attributive phrases, as in Hausa *mutane da yawa* (people with abundance) 'many people', or in possessive-like constructions, as in Akan *nnipa no nyinaa* (people the wholeness) 'all the people' (cf. *nnipa no ntade* (people the clothes) 'the people's clothes'). Another involves *verbs* of quantity, such as Akan *dɔɔso* 'be enough/much', as in:

(110)   Wɔnoaa   aduan a   ɛ*dɔɔso*
         they. cooked food   REL it. is. enough
         'They cooked enough / a lot of food'

It is, of course, very common for plurality to be expressed by affixes on nouns, whether by suffixation, as in English *houses*, *fingers*, or by prefixation, as in Ilocano *balbalay* 'houses', *ramramay* 'fingers' (where the plural prefix is a reduplication of the first three segments of the noun stem – cf. *balay* 'house', *ramay* 'finger'). Less common, but attested in certain synthetic languages, is the use of noun affixes to express other quantifier meanings: for example Yana *hanmau-* 'many', as in *hanmauyaa* 'many people'.

The next group of noun adjuncts to be considered is the *classifiers*. These are words which are required by the syntax of certain languages when a noun is also modified by a numeral. (In some languages, such as Mandarin Chinese, classifiers are also required when nouns are modified by demonstratives, or by one of certain non-numerical quantifiers. In Thai, on the other hand, classifiers are obligatory only with a subset of numerals, those expressing 'small definite numbers' – see Adams and Conklin (1974).) The closest English analogue to classifiers are the words that follow the numerals in expressions like *two heads of lettuce* or *three ears of corn*. But while in English only a relatively small group of nouns are not directly modified by numerals, in languages with classifiers this is true of all nouns. Thus in English one says *two boys*, *three dogs*, *four houses*, etc. But in Thai the equivalent expressions must all have classifiers: *deg sɔɔŋ khon* (boy two classifier) 'two boys', *maa saam tua* (dog three classifier) 'three dogs', *baan sii laŋ* (house four classifier) 'four houses', etc.

The number of classifiers found in a language may be quite large. Thus Warotamasikkhadit (1972) lists over sixty classifiers that occur in Thai, and acknowledges that the listing is incomplete. In some cases a given noun may co-occur with one of two or more different classifiers, in which case each classifier usually has a distinct meaning. Thus in the Thai examples *kluay sii kɔɔ* (banana four classifier) 'four banana trees (in a cluster)', *kluay sii wii*

(banana four classifier) 'four bunches of bananas', *kluay sii bai* (banana four classifier) 'four bananas', the meaning difference is obviously conveyed by the distinct classifiers.

As the examples from Thai have suggested, the classifier or classifiers that may occur with a given noun are selected by that noun. Thus the classifier *khon* is selected by *deg* 'boy' while the classifier *tua* is selected by *maa 'dog'*. But although the selection of classifiers is in part semantically based (thus *khon* is used only for humans), there is not always any obvious semantic basis for the selection of a particular classifier by a particular noun, and 'sometimes native speakers themselves are not sure which classifier is to be used in agreement with a certain noun' (Warotamasikkhadit (1972:23)). Evidently the situation is rather similar to that found in the inflectional grammatical-gender systems of Indo-European or Bantu languages, where there are some generalizations that can be made about the semantic correlates of the genders, but where there are also many cases in which the gender classification appears to be semantically arbitrary (see section 1.1, above).

The last group of noun adjuncts to be considered is the group of *articles*. Under this heading we wish to include, in addition to the words usually identified as definite and indefinite articles (e.g. English *the*, *a*), words that are sometimes identified as demonstrative adjectives or modifiers (e.g. *this* in *this man*, *that* in *that woman*). The reasons for grouping the demonstrative modifiers together with the (other) articles are both syntactic and semantic. Syntactically, demonstrative and other articles usually constitute a single distributional class, occurring in the same position in relation to the noun and other elements of the noun phrase, and not co-occurring in a single noun phrase: compare *a small woman*, *this small woman*; Akan *ɔbea ketewa bi* (woman small a) 'a small woman', *ɔbea ketewa yi* (woman small this) 'this small woman'. (There are exceptions, however. For example, in Hebrew the equivalent of the definite article is a prefix, and this prefix can co-occur with a demonstrative: e.g. *ha-ish ha-ze* (the-man the-this) 'this man'.) And semantically, demonstrative modifiers are like definite articles in being reference indicators. (Thus *this* often indicates that the referent of the following noun is close at hand, just as *the* often indicates that the referent of the following noun is assumed to have already been established.) In a good many languages, in fact, there are single words which may be translated 'the' or 'that' according to the context: for example German *die Frau* (the/that woman) 'the/that woman'; Akan *ɔbea no* (woman the/that) 'the/that woman'. (While the demonstrative modifiers are here grouped with the articles, for the reasons just indicated, it is also true that they usually have a close relation, both semantically and morphologically, to the demonstrative *pronouns* discussed in section 2.1.)

Articles may or may not show agreement with the nouns they modify. In Akan, for example, although nouns and (certain) adjectives are inflected

for number, the definite article and demonstrative *no* is invariable: cf. *ɔbea ketewa no* (woman small the/that) 'the/that small woman', *mmea nketewa no* (women plural-small the/that) 'the/those small women'. In German, on the other hand, the definite article and demonstrative *der/die*, etc., varies in form with the number, gender, and case of the noun it modifies: *der Mann* (the/that-nominative-singular-masculine man) 'the/that man', *die Frau* (the/that-nominative-singular-feminine woman) 'the/that woman', *das Buch* (the/that-nominative-singular-neuter book) 'the/that book', etc.

Languages that do not have articles may express equivalent meaning morphologically. For example, in Yuma, the demonstrative suffixes *-va*, *-nʸ*, and *-sa* are placed between the noun stem and the case marker: e.g. *ʔa·ve-va-c* (snake-this-nominative) 'this snake', *ʔa·ve-nʸ-c* (snake-that-nominative) 'that snake', *ʔa·ve-sa-c* (snake-that(distant)-nominative) 'that (distant) snake'. Similarly, in Tonkawa, a meaning of definiteness is expressed by a suffix *-ʔa·* before the case suffix on a noun, while a meaning of indefiniteness is expressed by the lack of this *-ʔa·*: for example *kʷa·n-ʔa·-la* (woman-the-nominative) 'the woman', *kʷa·n-la* (woman-nominative) 'a woman'.

The morphological indication of definiteness may be tonal rather than affixal. Thus in Bambara definiteness is expressed by a low final tone on the noun: for example *káfê* 'the coffee' (cf. *káfé* 'coffee' – the falling tone at the end of *káfê* results from the addition of a low final tone to an inherently high tone). There are also languages in which the definite–indefinite distinction is, in some instances at least, expressed by the case system. Thus in Southern Lappish (see Wickman (1955)) a definite direct object is in the accusative case while an indefinite direct object is in the nominative case, as in the following example:

(111)   a. jüktie treæwgəjdʹə dojtəmə
            when   skis(ACC)   one. has. made
            'when one has made the skis'

        b. jüktie treæwgah dajtθjh
            when   skis(NOM) they. make
            'when they make skis'

2.3   *Verb adjuncts*

This section is concerned with two classes of words that form phrasal constituents with verbs: *auxiliaries* and *verbal particles*. (The label *auxiliaries* seems preferable to the perhaps more common *auxiliary verbs* from a cross-linguistic point of view, since it allows for the inclusion of non-verbs in the class. While most auxiliaries are probably derived from verbs historically, and many can reasonably be identified as a subclass of verbs synchronically, there

are also cases – such as the Hausa examples cited below – where a synchronic analysis of auxiliaries as verbs seems questionable.)

Auxiliaries are words that express the tense, aspect, mood, voice, or polarity of the verb with which they are associated: i.e., the same categorizations of the verb as may be expressed by means of affixes (cf. section 1.2). English examples of auxiliaries expressing tense, aspect, and mood (respectively, future, perfect, and conditional) are:

(112)   John *will* understand
         John *has* understood
         John *would* understand

English also offers examples like the following of auxiliaries expressing voice (passive) and polarity (negative), in combination with tense:

(113)   John *was* understood
         John *won't* understand

Some representative examples from other languages are:
Bambara

(114)   a. U   *ye*   a   san
            they PAST.AFFIRM it buy
            'They bought it'

         b. U   *ma*   a   san
            they PAST.NEG   it buy
            'They didn't buy it'

         c. U   *bɛ*   a   san
            they PROG.AFFIRM it buy
            'They are buying it'

         d. U   *lɛ*   a   san
            they PROG.NEG   it buy
            'They aren't buying it'

Hausa (in this language a subject-pronoun morpheme and an auxiliary morpheme regularly combine to form a single word, the ordering of the two morphemes varying with the auxiliary used):

(115)   a. *Za*-ta   tafi
            FUT-she   go
            'She will go'

         b. Ta-*kan*   tafi
            she-HABIT go
            'She goes'

    c. Ta-*na*   tafiya
        she-PROG going
        'She is going'

    d. *Ba*-ta   tafiya
        PROG.NEG-she going
        'She isn't going'

And *Kannada* (examples from Upadhyaya and Krishnamurthy (1972)):

(116)   a. Baritta *iddiini*
           writing I.PROG
           'I am writing'

           b. Baritta *irtiini*
           writing I.PROG.HABIT
           'I am writing'

    In some languages sequences of two or more auxiliaries are allowed, in which case their order in relation to one another is generally fixed, as in the following examples from English and Tera (the latter from Newman (1970)):

(117)   John *must have been* sleeping

(118)   Ali *kə*   *ka*   *ɗa*   nji *ɫu*
        Ali SJNCT HABIT DISTANT eat meat
        'Ali should regularly eat meat (there)'

In other languages, such as Bambara (cf. (114)), only one auxiliary may occur in each clause.

    Greenberg (1963) has noted a correlation between the position of an inflected auxiliary in relation to the verb and other word order properties of the language. Stated in general terms, this correlation is to the effect that the position of an inflected auxiliary in relation to the verb will generally be the same as the position of the verb in relation to an object. Thus in languages where the verb precedes the object (e.g. English), an inflected auxiliary generally precedes the verb, while in languages where the verb follows the object (e.g. Kannada), an inflected auxiliary generally follows the verb. (Greenberg's own generalization is somewhat narrower than this: namely, 'In languages with dominant order VSO, an inflected auxiliary always precedes the main verb. In languages with dominant order SOV, an inflected auxiliary always follows the main verb' (p. 67). Greenberg thus does not propose a generalization about SVO languages. The data that he cites, however, show that, in almost all cases, inflected auxiliaries precede the verb in SVO, as in VSO, languages.)

    It should be noted that Greenberg's word order generalization applies only to *inflected* auxiliaries, so examples like those in (114) do not constitute an

exception to it. (That is, while Bambara is an SOV language in which the auxiliary precedes the main verb, Greenberg's generalization is irrelevant to Bambara since the auxiliary is uninflected in this language.) There is, however, one well-known systematic exception to Greenberg's word order generalization, having to do with so-called verb-second languages, such as German. These are languages which show SOV order in subordinate clauses, but in main clauses place a tense-bearing verb immediately after the initial constituent: cf. the position of the non-auxiliary verb *hat* 'has' in the following German examples:

(119) a. Ich weiss, dass er zu viel   Arbeit *hat*
          I   know that he too much work   has
          'I know that he has too much work'

     b. Er *hat* zu viel   Arbeit
        he has too much work
        'He has too much work'

     c. Heute *hat* er zu viel   Arbeit
        today has he too much work
        'Today he has too much work'

In such languages, auxiliaries follow the verb in subordinate clauses, thus conforming to Greenberg's generalization. In main clauses, however, a tense-bearing auxiliary immediately follows the initial constituent and thus precedes the non-tense-bearing verb. Note the position of the auxiliary *wird* 'will' in the following examples:

(120) a. Ich weiss, dass er zu viel   Arbeit haben *wird*
          I   know that he too much work   have   will
          'I know that he will have too much work'

     b. Er *wird* zu viel   Arbeit haben
        he will too much work   have
        'He will have too much work'

     c. Heute *wird* er zu viel   Arbeit haben
        today will he too much work   have
        'Today he will have too much work'

In languages that lack auxiliaries, the equivalent meanings are often expressed by verbal affixes. Some examples from Tagalog, all involving the verb base *luto* 'cook', are: *magluluto* 'will cook', *nagluluto* 'is cooking', *nagluto* 'has cooked', *nakakapagluto* 'can cook', *niluto* 'was cooked'. It is also possible, of course, for a language that does have auxiliaries to have such verbal affixes as well. Thus English uses an auxiliary to express the future tense (*will cook*) and an affix to express the past tense (*cooked*). Negation, like the other categorizations that may be expressed by auxiliaries, may be expressed affixally (as in Akan *n-kɔ*

(negative-go) 'doesn't go' – cf. *kɔ* 'goes'), but is also very commonly expressed by a distinct parts-of-speech class of *negators* (cf. section 2.5, below).

The second class of verb adjuncts to be discussed, the verbal particles, is a closed class of uninflected words that co-occur with certain verbs. Examples from English are the italicized words in:

(121)   John woke *up*, turned *off* the alarm, and switched the light *on*

In some cases the verbal particles may have clearly distinguishable locative or directional meanings, for example:

(122)   John kept his head *up/down*
         John carried the package *in/out*

In other cases, however, the particle forms an idiomatic lexical unit with the verb and does not carry any separable meaning: for example *up* in *wake up*, *give up*, *hurry up*; *down* in *break down*, *calm down*, *write down*.

Some examples of particles from languages other than English (respectively, German, Akan, and Ga'anda – the last from Paul Newman (1971)) are:

(123)   a. Er stand sehr früh   *auf*
            he stood very early on
            'He got up very early'

         b. Er sah den Wagen *an*
            he saw the car   at
            'He looked at the car'

(124)   a. Kofi gyee   Kwame *so*
            Kofi received Kwame on
            'Kofi answered Kwame'

         b. Kofi daa Kwame *ase*
            Kofi lay Kwame under
            'Kofi thanked Kwame'

(125)   a. ə   ʃənnda   wanɓəɓa   *in*
            PERF they. send medicine   along
            'They sent medicine along'

         b. Ni   xiy pirsh *kaɗə*
            I.FUT buy horse away
            'I will sell a horse'

In some languages, some or all of the verbal particles also occur as (and are historically derived from) adpositions (see section 2.2). Thus English *up* and *down* and German *auf* and *an* also occur as prepositions, while Akan *so* and *ase*

also occur as postpositions. In other languages, however, for example Ga'anda, the verbal particles are entirely distinct from adpositions.

While the particles are, in general, selected by the particular verbs with which they occur, and while they often join with the verbs in forming idiomatic units, the verbs and particles are, as the examples cited above have shown, not necessarily adjacent to one another. The syntactic rules of a language may specify that a verb and an associated particle may or must be separated from one another under certain circumstances. In English, for example, the object of a transitive verb may in most cases come between the verb and the particle, as well as after the particle (but cf. (53) where only one order is allowed with direct-object personal pronouns):

(126)　　John *looked* two words *up*
　　　　　John *looked up* two words

In German, if the verb is in clause-final position, the particle is prefixed to it, as in:

(127)　　Ich weiss, dass er sehr früh　*auf-stand*
　　　　　I　　know　that　he　very　early　up-stood
　　　　　'I know that he got up very early'

But if the verb follows the initial constituent of the clause (as it regularly does if the clause is a main clause and the verb is tense-marked – see above), the particle follows the verb, and is separated from it by other constituents of the verb phrase, as in (123). Similarly in Akan and Ga'anda a transitive verb and a particle are separated by the object of the verb (see (124) and (125)).

2.4　　Conjunctions

Conjunctions are words that are used to connect words, phrases, or clauses. Two general classes of conjunctions, *coordinating* and *subordinating*, are traditionally distinguished. The coordinating conjunctions are those that assign equal rank to the conjoined elements. (English examples are *and*, *or*, and *but*.) The subordinating conjunctions are those that assign unequal rank to the conjoined elements, marking one of them as subordinate to the other. (English examples are *whether*, *that*, *although*, etc.) These two classes of conjunctions are discussed in turn below.

Coordinating conjunctions generally occur *between* the elements that they conjoin. There is often evidence, however, that the conjunctions are more closely associated structurally with one of the conjuncts than with the other. One type of evidence to this effect is phonological, having to do with the points within a conjoined structure at which a pause (often reflected in writing by a comma)

is possible. In some languages, such as English, there is a potential for pause immediately *before* a coordinating conjunction but not immediately *after* one, while in others, such as Japanese, just the opposite is true – compare (128) from English and (129) from Japanese:

(128)   John, (and) Bill, and Tom came
           *John and, Bill and, Tom came

(129)   John to, Bill to, Tom ga   kita
           John and Bill and Tom SUBJ came
           'John, Bill, and Tom came'
           (cf. *John, to Bill, to Tom ga kita)

Thus in languages like English, coordinating conjunctions can be characterized as *prepositional*, since they form structural units with the conjuncts they precede, while in languages like Japanese they can be characterized as *postpositional*, since they form structural units with the conjuncts they follow.

It appears that the prepositional or postpositional character of the coordinating conjunctions that occur in a language are quite systematically associated with the language's general word order characteristics. Specifically, non-verb-final languages generally have the prepositional type of conjunction, verb-final languages the postpositional type. Further evidence to this effect is to be found in the positions of correlative – or paired – coordinating conjunctions, such as English *both–and* and *either–or*. In non-verb-final languages, such as English (see (130)) and Hausa (see (131)), correlative conjunctions typically precede each of the conjuncts, while in verb-final languages, such as Japanese (see (132)) and Turkish (see (133)), they typically *follow* each of the conjuncts:

(130)   *Both* John *and* Bill like Mary

(131)   *Da* Audu *da* Bello sun   ci   abinci
           and Audu and Bello they.PERF eat food
           'Both Audu and Bello have eaten'

(132)   Michiko *to* Michika *to* ga   gakusei desu
           Michiko and Michika and SUBJ student are
           'Both Michiko and Michika are students'

(133)   Šapkanɨ   *da* paltonu   *da* giy
           your. hat and your. coat and wear
           'Wear both your hat and your coat'

(As examples (131–3) suggest, correlative coordinate conjunction in most languages involves repeating the *same* conjunction, whether before each conjunct,

as in Hausa, or after each, as in Japanese and Turkish. Correlative conjunction in English is thus somewhat atypical.)

Languages may vary quite markedly in the types of constituents that they allow to be connected by means of coordinating conjunctions. In English, a very wide range of constituents may be connected in this way: nouns and noun phrases, verbs and verb phrases, adjectives, adverbs, prepositions, clauses, etc. In a good many other languages, on the other hand, coordinating conjunctions (or at least those that are the translation equivalents of *and*) are used primarily, or exclusively, to connect nouns and noun phrases. This is true of Hausa and Japanese, for example (although these languages do have ways of expressing the *semantic* equivalent of verb phrase conjunction, etc. – see below).

In this connection, it is interesting to note that in many languages 'and' and 'with' are expressed by the same word, as in the following Hausa and Japanese examples:

(134)   a. John *da* Bill sun   zo
          John and Bill they.PERF come
          'John and Bill came'

        b. John ya   zo   *da* Bill
          John he.PERF come with Bill
          'John came with Bill'

(135)   a. John *to* Bill ga   kita
          John and Bill SUBJ came
          'John and Bill came'

        b. John ga   Bill *to* kita
          John SUBJ Bill with came
          'John came with Bill'

If, as such examples suggest, the 'and' conjunction in these languages has developed historically from a prepositional or postpositional noun adjunct (cf. section 2.2), it is not surprising to find that it is used primarily for conjoining nominals.

Let us now consider some of the alternatives to coordinating conjunctions that languages may use to express the semantic equivalent. One such alternative is simple concatenation of the conjuncts, as in the following examples of verb-phrase coordination from Akan and Hausa respectively:

(136)   Nnipa no dii nam nomm bia
        people the ate meat drank beer
        'The people ate meat and drank beer'

(137)   Audu ya   tafi ofishinsa ya   yi aiki
         Audu he.PERF go   office his he.PERF do work
         'Audu went to his office and worked'

While such concatenative constructions are especially common for conjoining
verbs and verb phrases, they are by no means restricted to this function. There
are, for example, a good many languages that use concatenation for noun-phrase
coordination as well, either as an alternative to coordinating conjunctions (as
in Japanese and Turkish) or as the sole coordination strategy (as in Lahu – cf.
Matisoff (1973)).

   Some other coordination strategies that do not involve conjunctions are illus-
trated by the following examples, from Akan and Japanese respectively:

(138)   Yɛ-ne   wɔn abom   bio
         we-be. with them have. united again
         'We and they have united again'

(139)   John wa asa   okite,   kao o   aratta
         John TOP morning getting. up face OBJ washed
         'John got up in the morning and washed his face'

Example (138) (taken from Christaller (1875)) involves the coordination of
nominals. In this example the equivalent of *and* is expressed by the verb *ne*.
The fact that *ne* is properly analysed as a verb is clear from the forms of the
pronouns that precede and follow it: *yɛ-* is a form which occurs elsewhere
strictly as a *subject* pronoun prefixed to a verb, while *wɔn* is a form which
occurs elsewhere strictly as an *object* pronoun following a verb. The sentence
structure in (138) involves serial verbs – cf. Schachter (1974) – and may be
compared with that of a sentence like:

(140)   Yɛ-de   wɔn aba
         we-take them have. come
         'We have brought them'

Example (139) (from Kuno (1973)) involves verb-phrase coordination, and is
similar to the concatenative constructions cited above (e.g. (136–7)), except
that the first verb in (139) is a dependent form, the gerundive *okite* 'getting up'
(cf. *okita* 'got up').

   Let us turn now to the subordinating conjunctions. These are words that
serve to integrate a subordinate clause into some larger construction. Like their
coordinating counterparts, subordinating conjunctions may be prepositional
or postpositional, with the prepositional type common in non-verb-final lan-
guages, the postpositional type common in verb-final languages. The following

examples, from non-verb-final Tagalog and verb-final Uzbek respectively, are thus quite typical. (The subordinating conjunctions in these examples are *complementizers* – see below.)

(141)   Itinanong ko *kung* nasaan sila
          asked   I   COMP where they
          'I asked where they were'

(142)   Ula Hasan gayergæ ketkæn *dep*   suradɨ
          they Hasan where   went   COMP asked
          'They asked where Hasan had gone'

Three classes of subordinating conjunctions can be distinguished on the basis of their functions: *complementizers*, *relativizers*, and *adverbializers*. These are discussed below in turn.

Complementizers mark a clause as the complement of a verb (see (141–2)), noun (143), or adjective (144):

(143)   I question the claim *that* the earth is flat

(144)   I am afraid *that* I must leave

A good many languages have a complementizer that is rather transparently derived from the verb meaning 'say'. This is true, for example, of *dep* in (142), and it is also true of *sɛ* in the following example from Akan:

(145)   ɛyɛ nokware *sɛ*   mihuu no
          it. is truth   COMP I. saw him
          'It's true that I saw him'

(As (145) shows, complementizers that are derived from verbs meaning 'say' are by no means restricted to indirect quotation.)

One common alternative to the use of a complementizer is simply not to mark the subordinate status of a complement clause, as in (146) or its Hausa equivalent, (147):

(146)   He said it was raining

(147)   Ya   ce ana   ruwa
          he.PERF say there. is rain
          'He said it was raining'

Another alternative is to mark the subordinate status of the complement clause by nominalizing it: for example, by using a nominalized verb form and marking

the complement subject as a possessive, as in the following English and Uzbek examples:

(148)   John anticipated Mary's winning the prize

(149)   Ula Hasannɨ gayergæ ketkænini suradɨ
          they Hasan's where   his. going asked
          'They asked where Hasan had gone'

(The nominalization construction of (149) alternates with the complementizer-marked clause of (142).)

   Relativizers are markers of relative clauses. Two examples, from Hausa and Akan respectively, are:

(150)   Na   ga mutumin *da* ya   yi aikin
          I.PERF see the. man REL he.PERF do the. work
          'I saw the man who did the work'

(151)   Ɔbarima *a*   minim no te   hɔ
          man   REL I.know him lives there
          'A man whom I know lives there'

Note that relativizers are not the same as relative pronouns (which are discussed in section 2.1). Relativizers merely mark the clause in which they occur as relative, while relative pronouns in addition have some nominal function within the clause. If we compare the relativizers of (150) and (151) with the relative pronouns in their English translations, we can see that Hausa *da* and Akan *a* have no nominal function, while *who* and *whom* function as subject and object respectively of the relative clauses in which they occur.

   Languages that do not use relativizers to mark relative clauses may use relative pronouns or special relative verb forms, as in the Quechua example in (152) (from Weber (1976)), or may simply leave the relative clause unmarked, as in the Japanese example in (153) or its English translation:

(152)   Maqa-ma-q runa fiyu
          hit-me-REL man bad
          'The man who hits me is bad'

(153)   Kore wa watakusi ga   kaita hon desu
          this TOP I   SUBJ wrote book is
          'This is a book I have written'

   Adverbializers mark clauses as having some adverbial function, such as the expression of time, purpose, result, etc. (See Thompson and Longacre in vol. II, chapter 5, for a detailed typology of adverbial clauses.) In some languages, many of the words that serve as adverbializers also serve as prepositional or

postpositional noun adjuncts (cf. section 2.2), as the following English, Japanese, and Hausa examples illustrate:

(154) John left *after* $\begin{cases} \text{Sally arrived} \\ \text{the game} \end{cases}$

(155) Ressya ga   tuku
       Ohiru $\Big\}$ *made* kore o   site-kudasai

       train   SUBJ arrive
       noon $\Big\}$ until this OBJ do-please

       'Please do this until $\begin{cases} \text{the train arrives'} \\ \text{noon'} \end{cases}$

(156) *Da* $\begin{cases} \text{sun} & \text{zo} \\ \text{dare} \end{cases}$ zamu   yi rawa

       at $\begin{cases} \text{they. PERF come} \\ \text{night} \end{cases}$ we.FUT   do   dancing

       'As soon as they come $\Big\}$ we'll dance'
       'At night

In some languages, an adverbializer in a subordinate clause may be optionally paired with another conjunction occurring in the main clause. English, for example, can pair *if* and *then* as in:

(157) If John goes, (then) Bill will too

In Vietnamese, there are many such optional pairings, as in the following examples (from Binh (1971)):

(158) *Vi*   anh   mach thay giao, (*ma*)   Ba phai phat
       because older brother report teacher (therefore) Ba PASS punish
       'Because you reported him to the teacher, Ba was punished'

(159) *Khi* toi den, (*thi*) Ba di roi
       when I   arrive (then) Ba go already
       'When I arrived, Ba had already gone'

(160) *Tuy*   Ba noi nhanh (*nhung*) toi cung-van heiu   duoc
       although Ba speak fast   (but)   I   still   understand possible
       'Although Ba talked fast I could still understand him'

There are also cases where an adverbializer in a subordinate clause is *obligatorily* paired with a conjunction in the main clause. In counterfactual conditional sentences in Hausa, for example, the counterfactual conjunction *da* must appear

in both the antecedent clause and the consequent clause (at least if the latter has a perfect-aspect predicate):

(161)   *Da* an   tambaye su,   *da* sun   yarda
            CF one. PERF ask   them CF they. PERF agree
            'If they had been asked, they would have agreed'

Various alternatives to the use of adverbializers are discussed in Thompson and Longacre, vol. II, chapter 5. These include simple juxtaposition of clauses and the use of special subordinate verb forms. Additional examples of the latter are provided by the following Eskimo sentences (from Harper (1974)):

(162)   Qiu-ga-ma   isiqpunga
            cold-because-I I. am. coming. in
            'Because I'm cold, I am coming in'

(163)   Audla-ru-vit   quviasutjanngittunga
            go away-if/when-you I. will. be. unhappy
            'If/When you go away, I will be unhappy'

2.5   *Other closed classes*

This section surveys some of the more widespread closed parts-of-speech classes not discussed in previous sections. The classes to be surveyed are: *clitics*, *copulas* and *predicators*, *emphasis markers*, *existential markers*, *interjections*, *mood markers*, *negators*, and *politeness markers.* They are discussed below in the listed order.

Clitics are words that occur in a fixed position in relation to some other sentence element. (If the fixed position is before the other element, the clitics are sometimes called *proclitics*; if after, they are sometimes called *enclitics*.) In some languages, clitics regularly follow the first word of the clause in which they occur. This is true, for example, of Tagalog, as is illustrated by the position of the clitic *daw* 'they say' in the following sentences:

(164)   a. Darating   *daw*   si   Pedro bukas
               will. arrive they. say TOP Pedro tomorrow
               'They say Pedro will arrive tomorrow'

        b. Hindi *daw*   darating   si   Pedro bukas
               NEG   they. say will. arrive TOP Pedro tomorrow
               'They say Pedro won't arrive tomorrow'

        c. Bakit *daw*   hindi darating   si   Pedro bukas?
               why   they. say NEG will. arrive TOP Pedro tomorrow
               'Why do they say Pedro won't arrive tomorrow?'

(The placement of clitics in Tagalog is actually somewhat more complex than these examples suggest: see Schachter and Otanes (1972:429–35).) In other languages, clitics occupy a fixed position in relation to a verb. In French, for example, clitics immediately precede the verb that governs them, except that they follow an affirmative imperative. Note the position of the clitic *y* 'there' in the following French sentences:

(165) a. Elle *y*   reste aujourd'hui
           she there stays today
           'She is staying there today'

       b. Il faut   *y*   rester aujourd'hui
          it is.necessary there to.stay today
          'It's necessary to stay there today'

       c. Restez-*y* aujourd'hui
          stay-there today
          'Stay there today'

In addition to having a fixed position in relation to other sentence elements, clitics also generally have a fixed, or partly fixed, position in relation to one another. Thus the three clitics in the French example in (166) must occur in the order shown, while the four in the Tagalog example in (167) allow the order variation shown, but no other:

(166) Personne *ne   nous en*   donne
       no. one   NEG us   some gives
       'No one gives us any'

(167) Nagtatrabaho *ka   na* $\left\{ \begin{array}{ll} ba & daw \\ daw & ba \end{array} \right\}$ roon?

       are.working   you now $\left\{ \begin{array}{ll} \text{Q} & \text{they.say} \\ \text{they.say} & \text{Q} \end{array} \right\}$ there

       'Do they say you are working there now?'

Since the class of clitics is positionally defined, it may cut across parts-of-speech classes that are defined on a functional basis. For example, the class of clitics in French includes the negator *ne*, the object and reflexive pronouns, *y* 'there', and *en* 'from there, some, etc.', and the class of clitics in Tagalog is even more heterogeneous, as is clear from examples such as (167) and (168):

(168) Hindi *pa man lamang tuloy*   *siya* nakakapagalmusal
       NEG   yet even just   as.a.result he   can.have.breakfast
       'As a result, he hasn't even been able to have breakfast yet'

Given this kind of heterogeneity, there is little of a systematic character that can be said about the types of elements that are likely to show clitic behaviour, and thus about the probable counterparts of these elements in other languages. It may be noted, however, that clitics are likely to be phonologically light words, relatively short and/or unstressed, and that, cross-linguistically, personal pronouns (which are usually phonologically light) seem to show more of a tendency to cliticize than any other single type of element. (For further discussion of clitics, with emphasis upon their differentiation from affixes, see Bickel and Nichols, in vol. III, chapter 3, section 2.2.)

Copulas are words used to indicate the relation between a subject and a predicate nominal or adjective. Many languages use a subset of *verbs*, the copulative verbs (see section 1.2), to indicate this relationship. This is true, for example, of English, which has copulative verbs like *be*, *become*, etc. In other languages, however, the copulas are clearly not verbs, and have quite distinct grammatical properties. In Hausa, for example, verbs precede their objects and are inflected for tense–aspect. Copulas, on the other hand, follow the predicate nominal and are uninflected except for gender, as in the following examples:

(169)   a. Ita   yarinya *ce*
             she girl   COP
             'She is a girl'

           b. Shi yaro *ne*
             he   boy   COP
             'He is a boy'

(*Ce* is the copula used with feminine singulars, *ne* the one used in all other cases.)

In some languages a distinction is made between copulas and what may be called *predicators*. The latter are used to mark predicate nominals when there is no overt subject. In Bambara, for example, the predicator *don* is distinguished from the copulative verb *ye*:

(170)   Alamisadon *don*
           Thursday   PREDICATOR
           'It's Thursday'

(171)   Bi   ye   Alamisadon *ye*
           today PRES Thursday   be
           'Today is Thursday'

In other languages, however, the same words predicate nominals with and without subjects. Compare the following Hausa example with (169):

(172)   Audu *ne*
         Audu COP
         'It's Audu'

As was noted in section 1.2, there are languages that do not use copulas (or copulative verbs) to indicate the relation between a subject and a predicate nominal or adjective. In such languages the relation is indicated by juxtaposition, as in the following Ilocano examples:

(173)   a. Ina   daydyay babae
            mother that   woman
            'That woman is a mother'

         b. Napintas daydyay babae
            beautiful that   woman
            'That woman is beautiful'

There are also languages that use juxtaposition to express the relation in the *present*, but for non-present times use a tense-marked copulative verb, as in the following Swahili examples:

(174)   a. Hamisi mpishi
            Hamisi cook
            'Hamisi is a cook'

         b. Hamisi $\begin{Bmatrix} \text{alikuwa} \\ \text{atakuwa} \end{Bmatrix}$ mpishi

            'Hamisi $\begin{Bmatrix} \text{was} \\ \text{will be} \end{Bmatrix}$ a cook'

Emphasis markers are words that emphasize a predicate. (Words that emphasize a *nominal* are here called *contrast markers*, and are included in the category of *discourse markers* treated in section 2.2 – cf. example (108).) Examples from Vietnamese and Thai respectively, are:

(175)   Ong Ba *co*   xem quyen truyen ay
         Mr Ba EMPH read book   story that
         'Mr Ba *did* read that novel'

(176)   Naarii, kin Kaaw *sia*
         Nari   rice eat   EMPH
         'Nari, *do* eat your rice'

The usual English equivalent of an emphasis marker is a stressed auxiliary verb, as in the translations of (175) and (176), but in colloquial English *so* and *too*

are sometimes used as emphasis markers contradicting something that has been said or implied:

(177)   I am *so/too* telling the truth

Languages that do not have emphasis markers may be able to express the semantic equivalent by means of stress, and even in languages *with* emphasis markers this means may be available. Thus in Thai, according to Warotamasikkhadit (1972), the use of the emphasis marker *sia*, as in (176), is equivalent to placing emphatic stress on the verb.

   Existential markers are words which are equivalent to English *there is/are*, etc. Examples, from Hausa and Spanish respectively, are:

(178)   *Akwai* littafi a kan   tebur
          EXIST book at top. of table
          'There is a book on the table'

(179)   *Hay* muchos libros en la   biblioteca
          EXIST many   books in the library
          'There are many books in the library'

Some languages also have distinct negative existential markers, as in the following Hausa example:

(180)   *Babu*   littafi a kan   tebur
          EXIST.NEG book at top. of table
          'There isn't a book on the table'

   Languages that do not have existential markers often use verbs meaning 'be (located)' to express equivalent meanings, as in the following examples from Akan and Japanese:

(181)   Sika   bi   *wɔ*   me foto mu
          money some is.located my bag in
          'There is some money in my bag'

(182)   Yama   ni ki   ga   *aru*
          mountain on tree SUBJ is
          'There are trees on the mountain'

It is also quite common for there to be a close relation between existential and *possessive* constructions. For example, a word-by-word translation of the French existential idiom *il y a* is 'it there has', and in Tagalog the same words are used as existential and possessive markers:

(183)   a.   Mayroon-g ⎫
             Wala-ng   ⎭   libro   sa   mesa

             EXIST/POSS-LINK   ⎫
             EXIST/POSS(NEG)-LINK   ⎭ book   on   table

             'There { is / isn't } a book on the table'

        b.   Mayroon-g ⎫
             Wala-ng   ⎭   libro   ang   bata

             EXIST/POSS-LINK   ⎫
             EXIST/POSS(NEG)-LINK   ⎭ book   TOP   child

             'The child { was / doesn't have } a book'

Interjections are words that can constitute utterances in themselves, and that usually have no syntactic connection to any other words that may occur with them. English examples are *hello*, *ah*, *aha*, *bah*, *oh*, *wow*, etc. The class of interjections of a language often includes words which are phonologically distinctive. For example, English words must in general contain at least one vowel sound, but interjections like *hmm*, *pst*, and *shh* are vowelless. And in many languages clicks (sounds produced with a velaric air stream) can occur in interjections (as in English *tsk-tsk*), but not elsewhere. All interjections are deictic (see Comrie and Thompson, vol. III, chapter 6, on deixis).

Although there are a good many linguistic descriptions that fail to mention interjections, it seems likely that all languages do in fact have such a class of words. In the case of extinct languages interjections may not be attested in the written records because of the generally informal, colloquial character of this word class. In the case of modern languages, the omission of interjections from a linguistic description probably just signifies that the description is incomplete.

Mood markers are words that indicate the speaker's attitude, or that solicit the hearer's attitude, toward the event or condition expressed by a sentence. One common type of mood marker is the *request marker*, as exemplified by English *please*. Some others are illustrated by the Japanese examples of (184) (from Kuno (1973)), and the Tagalog examples of (185):

(184)   a.   Kore   wa   hon   desu   *yo*
             this   TOP book   is   STATEMENT
             '(I am telling you that) this is a book'

        b.   Kore   wa   hon   desu   *ka*?
             this   TOP book   is   Q
             'Is this a book?'

c. John   wa   baka   *sa*
John   TOP   foolish   STATEMENT
'(It goes without saying that) John is a fool'

(185)   a. Mabuti *a* ang ani?
good   Q TOP harvest
'Is the harvest good?'

b. Mabuti *kaya*   ang ani?
good   Q.SPECULATIVE TOP harvest
'Do you suppose the harvest will be good?'

c. Mabuti *sana* ang ani
good   WISH TOP harvest
'I hope the harvest is good'

(The Tagalog mood markers of (185) all also belong to the class of *clitics* – see above.) In languages that do not use mood markers, the semantic equivalent may be expressed in a wide variety of ways: for example by word order and intonation (as in English statements and questions), by verb inflections (cf. section 1.2) or auxiliary verbs (cf. section 2.3), or by explicit attitudinal expressions (e.g. *I hope*, *do you suppose*), etc.

   Negators are words like English *not*, which negate a sentence, clause, or other constituent. As was noted above, some languages have distinctive existential negators. In Tagalog, for example, the existential negator is *wala* (cf. (183)), while the general negator is *hindi* (cf. (168)). It is also quite common for languages to have distinctive *imperative/optative* negators: for example Tagalog *huwag*, as in (186).

(186)   a. *Huwag* kayo-ng   umalis
NEG   you-LINK   leave
'Don't leave'

b. *Huwag* siya-ng   pumarito
NEG   he-LINK   come. here
'He shouldn't come here'

   In some languages negation is regularly expressed by a *pair* of negative words. This is true, for example, of standard French, where negation requires the negative clitic *ne* plus some other negative word, as in (187).

(187)   a. Jean *ne*   veut   *pas* manger
Jean NEG wants not to. eat
'Jean doesn't want to eat'

b. Jean *ne*   veut   *rien*   manger
Jean NEG wants nothing to. eat
'Jean doesn't want to eat anything'

It is also true of general negation in Hausa, in which low-tone *bà* precedes and high-tone *bá* follows the constituent being negated, thus very neatly indicating the scope of the negation. For example:

(188)   a. *Bà*   Halima ta   yi *bá*
           NEG Halima she.PERF do NEG
           'It's not the case that Halima did it'

        b. Halima *bà*-ta   yi *bá*
           Halima NEG-she.PERF do NEG
           'Halima didn't do it'

        c. *Bà*   Halima *bá*   ta   yi
           NEG Halima NEG she.PERF do
           'It's not Halima who did it'

Languages that do not use negators may express negation by means of a verbal affix, as in Akan *ɔ-n-kɔ* (he-negative-go) 'he doesn't go' or Tonkawa *yakp-ape-n-o* (strike-negative-progressive-3rd person-present-declarative) 'he is not striking him'. There are also languages in which negation is expressed by an auxiliary verb – cf. section 2.3 for examples.

The last closed parts-of-speech class to be discussed is the class of *politeness markers*. These are words which are added to sentences to express a deferential attitude toward the person addressed. In Tagalog, for example, there are two politeness markers, *po* and *ho*, either of which may be added to any sentence the speaker wishes to render polite. (*Po* is more polite than *ho*; to borrow the terminology of Kuno (1973), *po* may be called 'superpolite'.) In some other languages, such as Japanese, the expression of politeness involves, instead of markers, a special polite vocabulary: for example *ee* 'yes (polite)', *hai* 'yes (superpolite)' vs *un* 'yes (informal)'; *boku* 'I (polite or informal)', *watakusi* 'I (polite or superpolite)' vs *ore* 'I (informal)'. (Japanese also has a special polite *affix*, *-mas-*, which is added to a verb in polite speech: e.g. *ake-mas-u* (open-polite-present) 'open (polite)' vs *ake-ru* (open-present) 'open (informal)'.) The use of special polite forms for 'you' is particularly common: for example Spanish *usted* 'you (polite-singular)', *ustedes* 'you (polite-plural)' vs *tú* 'you (informal-singular)', *vosotros* 'you (informal-plural)'.

This concludes our survey of closed parts-of-speech classes, as well as of parts-of-speech classes in general. While certain minor classes have been ignored, the great majority of the parts of speech encountered in the languages of the world have been covered, and on the basis of the material presented here, the field worker investigating an unfamiliar language should be reasonably well prepared for whatever parts-of-speech system he or she meets.

3   Suggestions for further reading

Linguistics textbooks generally make useful comments on parts of speech but do not present chapters on the subject. An excellent unified discussion can be found in Sapir (1921:116–19). Jespersen (1924) has three early chapters on parts of speech – 4, 5 and 6 – and he includes the distinctions he builds up in those chapters in subsequent chapters as he presents his point of view on language. We make this recommendation as we have found this point of view useful.

Some of the best treatments of the topic can be found in writings which are defending a point of view. Below are some works which will be accessible to those starting out on the topic. In addition to the specific writings we recommend, we include the volume edited by Vogel and Comrie, entirely on parts of speech. Some of that volume presents theoretical perspectives intended for people advanced in the field, but newly initiated readers can benefit from some of the chapters. When the writings concern just English, a more general cross-linguistic approach can be inferred: Bhat (2000); Bolinger (1967); Dixon (1991) – start with pages 6–9; Langacker (1987); Vogel and Comrie (2000); Wierzbicka (1986, 2000).

2 Word order

Matthew S. Dryer

0 Introduction

One of the primary ways in which languages differ from one another is in the order of constituents, or, as it is most commonly termed, their word order. When people refer to the word order of a language, they often are referring specifically to the order of subject, object, and verb with respect to each other, but word order refers more generally to the order of any set of elements, either at the clause level or within phrases, such as the order of elements within a noun phrase. When examining the word order of a language, there are two kinds of questions one can ask. The first question is simply that of what the order of elements is in the language. The second question is that of how the word order in the language conforms to cross-linguistic universals and tendencies. Our discussion in this chapter will interweave these two kinds of questions.

1 Some basic word order correlations

1.1 Verb-final languages

We will begin by examining a few of the word order characteristics of three verb-final languages, languages in which the verb normally follows the subject and object. Consider first Lezgian, a Nakh-Daghestanian language spoken in the Caucasus mountains, in an area straddling the border between Azerbaijan and Russia (Haspelmath (1993)). The example in (1) illustrates the verb-final order in Lezgian.

(1) Alfija-di   maq̃ala kx̂e-na
     Alfija-ERG article write-AORIST
     S   O   V
     'Alfija wrote an article'

The order in (1) is more specifically SOV (subject-object-verb), thus illustrating that not only do the subject and object both precede the verb, but the

subject (occurring in the ergative case) precedes the object as well, if both are overtly expressed. Most verb-final languages are SOV, though there are reported instances of languages which are OSV, and other verb-final languages in which there is considerable freedom in the order of subject with respect to object. SOV languages are the most widespread word order type among the languages of the world.

Lezgian has a number of other word order characteristics which are typical of verb-final languages. Among these is the fact that manner adverbs (Adv), like objects, precede the verb, as in (2).

(2)   Mirzebeg-a   k'ewi-z   haraj-na: '. . .'
        Mirzebeg-ERG   strong-ADV shout-AORIST
                       Adv   V
        'Mirzebeg shouted loudly: ". . ."'

Lezgian employs postpositions (Po), which follow the noun phrase they combine with, rather than prepositions (which would precede), as in (3).

(3)   duxtur-rin   patariw
        doctor-GEN.PL to
        NP   Po
        'to doctors'

Genitive noun phrases (G), noun phrases modifying a noun and expressing possession or a relationship like kinship, precede the noun, as in (4).

(4)   Farid-an   wax
        Farid-GEN   sister
        G   N
        'Farid's sister'

Another characteristic of Lezgian that is typical of verb-final languages is that in comparative constructions, the order is standard of comparison (St) followed by the marker of comparison (M) followed by the adjective, as in (5).

(5)   sad müküda-laj   žizwi   ask'an-zawa
        one other-SUPEREL   a.little low-IMPERF
             St   M   Adj
        'one is a little shorter than the other one'

The standard is a noun phrase to which something is being compared, in (5) the noun phrase *müküda* 'the other'. The adjective is *ask'an* 'low'. The marker is a morpheme combining with the standard and indicating that the standard is being compared with something; in (5), the marker is the superelative case suffix *-laj* on the noun *müküda* 'other'. The marker is realized in different languages

in various ways, by affixes or by separate words, and, if a separate word, by various parts of speech.

Finally, the example in (6) illustrates how adverbial subordinators, markers of adverbial subordinate clauses, occur at the end of the subordinate clause in Lezgian, as a suffix on the verb, illustrated here by the subordinator -*wiläj* 'because'.

(6)   ruš-az   reğü   x̂ana k'an tuš-ir   -wiläj
        girl-DAT ashamed be   want be.NEG-PTCPL-because
        Clause   Subord
        'because he did not want the girl to be embarrassed'

When we examine two other verb-final languages from different parts of the world, we find that they resemble Lezgian in each of the characteristics noted above. The first of these languages is Slave, an Athapaskan language spoken in northern Canada (Rice (1989)). As in Lezgian, the normal order is SOV, as in (7).

(7)   t'eere lį̱   ráreyį̱ht'u
        girl   dog 3.hit
        S   O   V
        'the girl hit the dog'

The examples in (8) illustrate how Slave resembles Lezgian in each of the other characteristics observed.

(8)   a. dzá dahehłe   b. dene hé
           bad   1.dance   man with
           Adv V   NP Po
           'I dance badly'   'with the man'

        c. ʔabá gok'eríʔeé
           father jacket
           G   N
           'father's jacket'

        d. sǫdee   nǫdee   ts'ę̱ʔóné hį̱shá
           1SG:older.brother 2SG:older.brother than   3.big
                                    St   M   Adj
           'my brother is bigger than your brother'

        e. [kǫ́   seghǫ húle]   t'áh
           match 1SG.for 3:be.none because
           Clause   Subord
           'because I had no matches'

Siroi, a Madang language spoken in Papua New Guinea is also SOV (Wells (1979)), as illustrated in (9).

(9)   fe-nge   tango make-te
       taipan-SPEC   man   bit-3SG.PRES
       S   O   V
       'a taipan bites a man'

The examples in (10) illustrate how it exhibits the same characteristics as Lezgian and Slave:

(10)   a. nu pitik   kin-it   b. mbanduwaŋ mbi
          he quickly go-3SG.PRES   bow   INSTR
             Adv   V   NP   Po
          'he is going quickly'   'with a bow'

       c. tisa   tuku age   d. [ne kuayar-at]   tukunu
          teacher of   dog   you steal-2SG.PAST   because
                   G   N   Clause   Subord
          'the teacher's dog'   'because you stole it'

(We do not have information on how comparative meanings are expressed in Siroi.)

1.2   *Verb-initial languages*

Let us turn now to three instances of verb-initial languages, languages in which the verb normally precedes both the subject and the object. Such languages are much less common than verb-final languages. What we will see is that these languages exhibit the opposite characteristics from those that we saw in the three verb-final languages discussed above. The first verb-initial language we will look at is Fijian, an Austronesian language spoken on the island of Fiji in the Pacific Ocean (Dixon (1988)). Both the subject and the object follow the verb in Fijian, though they can occur in either order with respect to each other. Thus, the sentence in (11) can be interpreted either as 'the old person saw the child' (VOS) or as 'the child saw the old person' (VSO), and both orders are common in usage.

(11)   e   rai-ca   a   gone a   qase
       3SG see-TRANS   ART child ART old.person
            V   S/O   S/O
       'the old person saw the child' or 'the child saw the old person'

Note that a third person singular clitic agreeing in person and number with the subject precedes the verb.

The following examples illustrate how Fijian has the reverse characteristics from those we have observed for verb-final languages. Manner adverbs follow the verb, rather than preceding:

(12)   bau   'ada va'a-totolo noo
           somewhat run ADV-quick ASP
                        V   Adv
       'try and run more quickly'

The language employs prepositions (Pr) rather than postpositions:

(13)   mai   Wairi'i
           from Wairi'i
           Pr   NP
       'from Wairi'i'

The genitive follows the possessed noun, rather than preceding.

(14)   a   liga-i   Jone
           ART hand-POSS John
                 N   G
       'John's hand'

Note that the possessed noun *liga* 'hand' in (14) bears a suffix *-i* indicating that it is possessed by someone. The order in comparative constructions is adjective-marker-standard, the opposite from what we saw in the verb-final languages:

(15)   e   vina'a ca'e o   Waitabu mai Suva
           3SG good   more ART Waitabu from Suva
                Adj   M   St
       'Waitabu is better than Suva'

Note that the subject intervenes between the adjective and the marker + standard in (15). And adverbial subordinators occur at the beginning of the subordinate clause, as in (16), where the first singular subject pronoun *u* of the subordinate clause cliticizes onto the subordinator *ni*.

(16)   ni=[u   sa   daga.daga va'a-levu]
           when=1SG ASP tired   ADV-great
           Subord   Clause
       'when I'm very tired'

The two other verb-initial languages we will examine resemble Fijian in exhibiting the opposite characteristics from those we saw in the three verb-final

languages. The first of these is Turkana, in the Nilotic subfamily of Nilo-Saharan and spoken in Kenya (Dimmendaal (1983)). Turkana is VSO, as in (17).

(17)   ὲ-sàk-ɪˋ   apaˋ   akìmʊj
        3-want-ASP   father.NOM   food
        V   S   O
        'father wants food'

(Most of the nouns in the Turkana examples cited in this chapter contain gender prefixes that are not indicated in the   glosses, since Dimmendaal does not gloss them, and it is not always clear what gender is involved. A number of other affixes in nouns and verbs are not glossed and are treated here as if they were part of the stems. Nominative case, used for subjects, is indicated by tone.)

The examples in (18) illustrate how other word order characteristics of Turkana are the same as those in Fijian.

(18)   a. ὲ-à-gʊ̀m-ɪˋ   nilèmʊ̤
            3-PAST-fire-ASP   blindly
            V   Adv
            'he fired blindly'

        b. ὲ-à-gʊ̀m-ɪˋ   à   atɔmɛ̤̀
            3-PAST-fire-ASP   PREP   gun
                               Pr   NP
            'he fired with a gun'

        c. itɔ̀ɔ̀   keŋˋ   à   [èdyaˋ   loˋ]
            mother   his   of   boy   this
            N   G
            'the mother of this boy'

        d. lògerɪ̤̀   [lo-e-putuk-ì-o   erotˋ]
            because   REL-3-muddy-ASP-VERB   road.NOM
            Subord   Clause
            'because the road is muddy'

The expression involving comparison involves the use of a verb meaning 'surpass' or 'supersede', as in (19).

(19)   ὲ-jɔkˋ   erotˋ   loˋ   ak-ìdwaŋ   ŋolˋ
        3-good   road.NOM   this INFIN-supercede   that
        'this road is better than that one'

While one might treat the verb *akìdwaŋ* 'supersede' as a marker, this is really the verb of a separate clause, and hence *akìdwaŋ ŋolˋ* is not modifying the adjective in the same way as marker plus standard in true comparative constructions. The

expression of equative comparison, however, employs the order AdjMSt, using a construction involving a single clause, as in (20).

(20)   a-wòs   ayɔ̀ŋ   à   ni-konì̜
         1SG-clever 1SG.NOM PREP LOC-your
         Adj   M   St
         'I am as clever as you'

   Lealao Chinantec, an Oto-Manguean language spoken in Mexico (Rupp (1989)), is also verb-initial, except that it is VOS rather than VSO:

(21)   ka^L-kiúʔ^M mi^{VH}-zïï^L-i   [za^M ni^M]
         PAST-strike.COMPL.3 CLSFR-head-1SG person that
         V   O   S
         'that person struck my head'

(The superscript capital letters in (21) indicate tones, which play a major grammatical role in Chinantec languages. The form of the verb *kiúʔ* ^{M}, including its tone, indicates that it is a transitive verb, with an inanimate object, that the aspect is completive and that the subject is third person. The low tone on the noun *zïï^L* (along with the suffix *-i*) indicates that its possessor is first person singular.)

   The following examples illustrate how Lealao Chinantec displays the same word order properties we have seen in Fijian and Turkana, and the opposite of what we saw in the three verb-final languages:

(22)   a. ʔi^H-u̜:^{LH}-i   ziú^L b. he:^{LH} nu:^{M}
            INTENTIVE-wash-1SG well   among weeds
            V   Adv   Pr   NP
            'I will wash it well'   'among the weeds'

       c. siá:^{VH} [diáʔ^L sïï^{M}iu̜:^{M}i]
           mother.3 PL   baby
           N   G
           'the mother of the babies'

       d. gá:^{M}i gïʔ^{VH} tï:^{VH}i   niu^M liaʔ^M sïï^M ké^L hniá^M
           big.3 more foot.2SG   2SG   like   as   of.1SG 1SG
           Adj   M   St
           'your foot is bigger than mine'

       e. kia:ʔ^{VH} [ʔa^{L}ʔe^M na^{M}fáʔ^{L}i   hɨ i:^{LH}-a^L]
           because not   significant pay-1SG
           Subord Clause
           'because my wages aren't sufficient'

1.3   SVO languages

Finally, consider three instances of SVO languages, which are neither verb-final nor verb-initial, since the subject precedes the verb while the object follows the verb. SVO languages are the second most widespread word order type among the languages of the world, more common than verb-initial, but less widespread than verb-final languages. What we will see is that these SVO languages strongly resemble the verb-initial languages rather than the verb-final languages with respect to the word order characteristics examined. Consider first English, which is SVO:

(23)   The woman saw the dog
              S   V   O

As in the verb-initial languages we examined, English employs prepositions:

(24)   on the table
          Pr NP

The order in comparative constructions is AdjMSt:

(25)   Nancy is more intelligent than Jeff
                            Adj   M   St

Note that the marker of comparison is the word *than*, rather than the word *more.* Most languages do not employ a word meaning 'more' in comparative constructions, using expressions that literally translate more like 'Nancy is intelligent than Jeff', although the marker of comparison in such languages might be considered to mean 'more than' rather than just 'than'.

   Adverbial subordinators occur at the beginning of the subordinate clause, as illustrated by the adverbial subordinator *because* in (26).

(26)   because it was raining
          Subord   Clause

In each of the above characteristics, English resembles the three verb-initial languages rather than the three verb-final languages.

   When we look at the order of genitive and noun in English, we find two constructions, one in which the genitive precedes the noun, as in (27a), the other in which the genitive follows the noun, as in (27b).

(27)   a. the box's cover
                 G   N
          b. the cover of the box
                 N   G

We discuss in section 2 below the general problem of dealing with cases in which both orders of a pair of elements occur in a language. The general strategy is to try to identify one of the two orders as in some sense more basic. We will assume that in the case of genitive and noun in English, neither order is basic relative to the other and that English should thus be classified as GN/NG, as a language in which both orders of genitive and noun occur and in which there are no strong arguments for treating one of these orders as basic. Note that the GN construction in (27a) employs the order typically associated with verb-final languages while the NG construction in (27b) employs the order associated with verb-initial languages.

In the case of manner adverbs, English again exhibits both orders, as in (28).

(28)   a. John slowly walked into the room
                      Adv   V

        b. John walked into the room slowly
                  V   Adv

Here, there are arguments that the order VAdv is the basic order. Among these arguments is the fact that, in other contexts, the order VAdv is strongly preferred, as illustrated by (29).

(29)   a. ?*John is slowly walking   b. John is walking slowly

If we can assume that the second order is basic, then we can say that English is VAdv.

In summary, we see that English resembles the verb-initial languages in all respects but one: it has both GN and NG word order for the order of genitive and noun. It turns out that English is not atypical as an SVO language in this respect: while in most SVO languages one order can be identified as basic, in some SVO languages the order we find is GN, while in others it is NG. The two other SVO languages we will look at also resemble verb-initial languages in their word order characteristics.

Hmong Njua, a Hmong-Mien language spoken in China (Harriehausen (1990)), is a second example of an SVO language:

(30)   Peter muab pob khuum rua Maria
        Peter give   gift   to   Maria
        S   V   O
        'Peter gave a gift to Maria'

The example in (30) also illustrates that Hmong Njua is prepositional, the preposition *rua* 'to' preceding its object. The examples in (31) illustrate how Hmong Njua resembles English and the verb-initial languages we examined in other word order characteristics.

(31) a. Moob   lab has lug txawv luag
         Hmoob red speak   strangely
                         V   Adv
         'Red Hmoob speaks strangely'

b. tsuv luj dula miv   c. lub   thawv saab sau
   tiger big than cat   CLSFR box   cover
         Adj M   St   G   N
   'a tiger is bigger than a cat'   'the box's cover'

The example in (31c) illustrates GN order in Hmong Ngua, like the construction in English *the man's hat* and like the verb-final languages we examined rather than the verb-initial languages, but, as noted above, this order is as common as NG order among SVO languages.

   As a third example of an SVO language, consider Tetelcingo Nahuatl, a Uto-Aztecan language spoken in Mexico (Tuggy (1977)):

(32) [sen-te   tlɔkatl] (Ø-)kɨ-pɪya-ya   [sen-te   puro]
     one-NUM man   he-it-have-IMPERF one-NUM burro
     S   V   O
     'a man had a donkey'

The examples in (33) illustrate how Tetelcingo Nahuatl resembles the verb-initial and the other two SVO languages we have examined, though we do not have an example with a manner adverb:

(33) a. i-pa   i-čɔ   b. i-čɔ   mali
        3SG-at his-home   3SG-home Mary
        Pr   NP   N   G
        'at his house'   'Mary's home'

c. yaha kačɪ wieyɪ ke   taha
   he   more big   than you
             Adj   M   St
   'he is bigger than you'

d. [kwɔk   walɔ-s]   nɪ-tla-cɪlini-s
   when   come-FUT   1SG-UNSPEC.OBJ-clang-FUT
   Subord Clause
   'when he comes, I will ring the bells'

   We see that, apart from the order of genitive and noun, SVO languages tend to be like verb-initial languages rather than like verb-final languages. Because SVO languages share with verb-initial languages the fact that the object follows the verb, we can say that it is the order of object and verb (rather than subject

and verb) that is crucial in predicting other word order characteristics. For this reason, it is common to refer to the two types of languages as OV languages and VO languages. We will see below that there are a variety of other characteristics that are predictable from whether a language is OV or VO, though in a few instances the order of subject and verb is relevant as well.

1.4   *Object-initial languages*

The discussion above illustrates the most common word orders, SOV, SVO, and verb-initial (which includes both VSO and VOS). The two remaining orders are OVS and OSV, both quite rare but both claimed to exist. The clearest example of an OVS language is Hixkaryana, a Carib language spoken in Brazil (Derbyshire (1979)), illustrated in (34).

(34)   toto   y-ahosɨ -ye   kamara
        man   3SUBJ.3OBJ-grab-REM.PAST   jaguar
        O   V   S
        'the jaguar grabbed the man'

While a number of languages have been claimed to be OSV, the evidence so far presented for these languages is less than convincing.

What word order characteristics are typical of object-initial languages? Unfortunately, the number of clear cases of such languages is sufficiently small that we cannot really answer this question with any confidence. The fact that the characteristics in other languages pattern with the order of object and verb would lead us to expect both OVS and OSV languages to pattern with SOV languages. In so far as we have evidence, this prediction seems to be true. For example, Hixkaryana is postpositional and GN, as illustrated in (35).

(35)   a. maryeya ke   b. Waraka kanawa-rɨ
            knife   with   Waraka canoe-POSSD
            NP   Po   G   N
            'with a knife'   'Waraka's canoe'

There are a number of languages in which the basic or most frequent order in transitive clauses containing a lexical subject and a lexical object is OVS, but in which the basic or most frequent order in intransitive clauses is SV. A clear case of such a language is Parï, a Nilotic language spoken in Sudan (Andersen (1988)):

(36)   a. ùbúr   á-pùot   dháag-ɛ̀
            Ubur   COMPLET-beat woman-ERG
            O   V   S
            'the woman beat Ubur'

b. dháagɔ á-mɪ́ɛl`
   woman COMPLET-dance
   S   V
   'the woman danced'

Characterizing such languages as OVS is somewhat misleading in that the word order really follows an ergative pattern Abs-V-(Erg). Note that Parï has an ergative case-marking system as well, with an overt ergative case marker illustrated on the subject in (36a) and a zero absolutive case.

1.5   *Interim summary*

We can summarize the patterns we have observed so far as follows:

SOV	*SVO*	*Verb-initial*
AdvV	VAdv	VAdv
NP + Po	Pr + NP	Pr + NP
GN	GN or NG	NG
StMAdj	AdjMSt	AdjMSt
ClauseSubord	SubordClause	SubordClause

As noted above, SVO and verb-initial languages pattern the same way, except for the order of genitive and noun: SVO languages are sometimes GN and sometimes NG, whereas verb-initial languages are generally NG. If we collapse SVO and verb-initial into VO and assume that OVS and OSV pattern with SOV, then the patterns can be described in terms of a contrast between OV and VO. Note that all of the characteristics we have discussed involve pairs of elements, except for the order in comparative constructions, where three elements are involved. However, the order in comparative constructions can be thought of as two pairs of elements, the order of standard and marker and the order of standard and adjective. There are in fact a few languages which show that these two pairs of elements need to be distinguished, where the order is MStAdj or AdjStM. For example, the order in Mandarin Chinese is MStAdj, as in (37).

(37)   Zhāngsan bī   tā   pàng
        Zhangsan COMPAR 3SG fat
                  M   St   Adj
        'Zhangsan is fatter than her/him'

Mandarin is MSt, the order associated with VO languages, but StAdj, the order associated with OV languages. Since the normal order in Mandarin is SVO, the StAdj order is atypical.

The question of what underlies these word order correlations is one on which there is an extensive literature (see references listed at the end of this chapter)

and is a topic that we will not discuss here. But it should be noted that a common view is that the characteristics associated with OV order are head-last or head-final, while those associated with VO order are head-first or head-initial. But Dryer (1992) argues that there are serious empirical problems with this view. For example, the notion that OV languages tend to be head-final and VO languages head-initial would lead us to expect modifiers of nouns to precede the noun in OV languages and follow the noun in VO languages. But as is discussed in section 7 below, this is not true for adjectives, demonstratives or numerals: none of these three elements correlates in order with the order of object and verb, preceding and following the noun with similar frequency in OV and VO languages. And articles exhibit the opposite correlation, preceding the noun more often in VO languages than they do in OV languages.

1.6   *Conclusion*

Before examining other word order characteristics, there are two general problems that we must address that arise in attempting to identify word order characteristics of a language. The first of these problems is that of identifying a basic order for two or more elements when more than one order exists in a language. The second problem is that of identifying instances of particular constructions in different languages. We discuss each of these topics in the next two sections.

2   **Identifying basic word order**

In most of the languages we have examined, we were able to classify the languages according to each of the various characteristics examined. In instances in which only one order of a pair of elements is possible in a language, this classification is straightforward. But many languages exhibit more than one order for at least some pairs of elements, and questions arise as to how to classify the language according to the characteristic in question. There is some variation in the practice of linguists on this question, both in terms of what criteria to employ in these instances and in terms of whether to classify a language at all when the criteria do not yield an obvious answer. These issues have been most widely discussed in the context of identifying a basic order of subject, object, and verb, but they apply to all pairs of elements.

One of the criteria that have been appealed to in such instances is that of frequency of usage. Considering first an extreme example, English allows OV order, as in *Paul, I like*, but this order is quite obviously much less frequent than the order VO. Where languages allow alternative orders, one order is often overwhelmingly more frequent. But in other instances, the differences in frequency may be much less extreme. For example, D. L. Payne (1990) reports that in a count of texts in Yagua, a language spoken in Peru, the order SV

occurred 114 times while the order VS occurred 257 times. In this case, both orders are relatively frequent, but VS outnumbers SV by just over 2 to 1. Is this a valid reason for considering VS basic? Linguists answer this question in different ways. A number of reasons have been offered for not treating VS as basic in such instances. One argument is simply that frequency should not be used as a criterion because it is not part of the grammar of the language. Another possible argument is that such differences in frequency might be an artefact of a particular set of texts, and that one might find very different frequencies in a different set of texts. A further argument is that even if the set of texts can be considered sufficiently varied for the difference in frequency to be considered typical, the fact that both orders are relatively common is more important than the fact that one order happens to be more common than the other.

In defence of frequency, it can be argued that differences in frequency often provide a more reliable test than other tests in that, where the difference is large enough, it will be intuitively obvious to the linguist working on the language, and often to speakers of the language as well, that one order is the 'normal' order. And frequency is a clear operational test; if one order is consistently more common across large enough samples of texts, then anybody examining such texts will arrive at the same conclusion. Finally, many of the conclusions in word order typology are based on grammatical descriptions in which there is flexibility of word order but in which one order is described as normal. In other words, in practice, frequency has been the primary criterion in word order typology. Furthermore, the universal tendencies associated with OV versus VO order are found in languages in which there is considerable flexibility of word order, even among languages in which one order outnumbers the other by a frequency of only 2 to 1. It should be noted, however, that frequency counts of some languages do not reveal one order as noticeably more frequent than the other. In the Auk dialect of Tlingit, for example, a text count (Dryer (1985)) for the order of subject and verb revealed VS outnumbering SV by 177 to 156. In a case like this, the difference in frequency is sufficiently small for it not to seem reasonable to say that VS is more frequent than SV, or that VS is basic.

A number of criteria other than relative frequency have been appealed to in determining basic order. If one order is in some way more restricted in its distribution, then that can be used as an argument that the other order is basic. An example is the argument in section 1.3 above for treating VAdv order as basic in English because there are environments in which the order AdvV is not used (*?*John is slowly walking*). The restriction in distribution might be over syntactic contexts, as in the preceding example, or it might be over lexical items. In Korowai (Van Enk and De Vries (1997)), an Awju language spoken in Irian Jaya on New Guinea, all adjectives can precede the noun, as in (38a), but a few, like the one meaning 'big', can also follow the noun, as in (38b).

(38)   a. lembul   nggulun
            bad   teacher
            'a bad teacher'

        b. yanop   khonggél-khayan
            man   big-very
            'a very big person'

We can say that AdjN order is basic because it has a less restricted distribution.

   Some languages have both prepositions and postpositions, but there are often more of one than the other. For example, in Taba (Bowden (1997)), an Austronesian language of Halmahera in Indonesia, there are five prepositions and one postposition. The example in (39a) illustrates one of the five prepositions (*ada* 'with'), the example in (39b) the one postposition (*li* 'locative').

(39)   a. n-pun   bobay   ada   ni   sandal do
            3SG-kill   mosquito   with 3SG.POSS   sandal REALIS
            'he killed the mosquito with his sandal'

        b. n-battalon   kurusi   li
            3SG-sit   chair   LOC
            'he's sitting on the chair'

Because there is only one postposition in the language, we can say that Taba is basically a prepositional language, and hence (since it is SVO) that it conforms to the expectation of an SVO language being prepositional. Having even one postposition is somewhat unexpected of an SVO language, though not as unusual as SVO languages that are basically postpositional. Even English has a few words that can be analysed as postpositions, such as *ago* (as in *three years ago*) and *notwithstanding* (as in *I have decided to run for re-election, my family's opposition notwithstanding*).

   A further distributional criterion is based on simplicity. In English, adjective phrases sometimes precede the noun, as in (40), and sometimes follow the noun, as in (41).

(40)   a. the tall woman
        b. the very tall woman

(41)   a. the woman taller than John
        b. the woman angry at John

Nor can these be reversed; the adjective phrases in (40) cannot follow the noun (*the woman tall*, *the woman very tall*) nor can the adjective phrases in (41) precede the noun (*the taller than John woman*, *the angry at John woman*). Hence we cannot use a distributional test based on one position being more restricted than the other. But the adjective phrases that follow the noun in (41)

are clearly more complex: they contain entire phrasal modifiers of the adjective and these phrasal modifiers can easily be rendered more complex, as in (42).

(42)   the woman taller than the man who John was talking to

In contrast, the simplest adjective phrases, those consisting of just the adjective, must precede the noun, as in (40a). By this criterion, we can say that the basic order of adjective phrase and noun in English is for the adjective phrase to precede the noun.

    A third type of criterion, beyond frequency and distributional criteria, is one based on pragmatics. It can often be argued that one order in a language is pragmatically neutral while the other has some added pragmatic effect. In English, for example, the OV order in (43a) and the VS order in (44a) both apparently add a special effect that is absent in the neutral orders in (43b) and (44b).

(43)   a. Mary, I saw
         b. I saw Mary

(44)   a. Into the room came the Prime Minister
         b. The Prime Minister came into the room

In Ilocano, an Austronesian language spoken in the Philippines, adjectives can either precede or follow the noun, but postnominal position is contrastive (Rubino (1998:40)). The more neutral order is given in (45a), the more contrastive order in (45b). (The difference in the form of the linking morpheme *(ng)a* in (45) simply reflects a phonologically conditioned alternation: *nga* before vowels, *a* before consonants.)

(45)   a. ti   nalaíng nga   ubíng
            ART smart   LINK child
            'the smart child'

         b. ti   ubíng a   nalaíng
            ART child LINK smart
            'the smart child (as opposed to the others)'

    It is often, however, not obvious that one order involves adding an additional element of meaning, as opposed to the two orders simply having a difference in meaning. For example, in Papago, a Uto-Aztecan language spoken along the US–Mexican border, OV order is associated with indefinite objects while VO order is associated with definite ones (D. L. Payne (1987)). It does not seem right to say that VO order involves the addition of definiteness or that OV order involves the addition of indefiniteness, so in this case there is little basis for describing one order as pragmatically neutral.

Descriptions of languages often describe an order in which an element occurs at the beginning of a sentence as involving topicalization, but it is often extremely difficult to give objective criteria for identifying the actual pragmatic effect of the topicalization, and it is often not clear that the label is being used in anything more than a syntactic sense, to say that an element is in initial position in the clause. In practice, it is usually difficult to justify claims that one order is pragmatically nonneutral, except in cases like OV and VS order in English, where other criteria point to the non-basic nature of these orders.

Over the history of generative grammar, various arguments have been offered for some order being the underlying or deep structure order. Often, these arguments are based on the overall grammar being somewhat simpler if one order is treated as the underlying order. The arguments often depend on the assumptions of a particular version of generative grammar at a particular point in time and no longer apply under later assumptions. And even under a given set of assumptions, there are often competing arguments for which order is basic. And while the notion of underlying order is sometimes assumed to be the same as basic order, and hence the arguments for one order being underlying are treated as arguments for that order being basic, it is not at all clear that the notions are the same.

In cases in which there is some doubt as to what order of a pair of elements in a language might be called basic, for example when different criteria conflict, it is probably best not to force the language into one category or another, but simply to classify it as a language in which neither order is clearly basic. And when there is such doubt, what is most important in describing a language is not the determination of the basic order, but the more detailed facts that lead to there being some doubt.

This chapter cites examples from a large number of languages and identifies one order as basic, usually without further discussion or elaboration as to what criteria were used. In many of these cases, either the word order is rigid or there seems to be little question as to which order is basic, regardless of one's criteria. In some cases, however, this may not be so. Most of the characterizations of languages in this chapter are based on characterizations in grammatical descriptions of the language and we follow the grammarians' characterization of orders, though this may mean in some cases that different criteria are assumed for different languages. In practice, this means that frequency is treated as the major criterion, since grammars most often contain descriptions like 'the normal order is for the adjective to precede the noun', and rarely do grammars discuss other possible criteria. Since, in most cases, frequency coincides with other criteria, this means that most characterizations in this chapter are consistent with other criteria as well.

It should be stressed that it is not clear that issues of what order is basic are relevant to actually describing languages, as opposed to deciding whether

the language conforms to cross-linguistic expectations. One can describe Taba as having five prepositions and one postposition and there is no need for any further comment on prepositions being basic. Similarly, one can describe the position of attributive adjectives in Ilocano as normally prenominal with post-nominal position contrastive, without also including in one's description the idea that AdjN order is basic. The question of whether these orders are basic only arises if one wants to ask whether the language conforms to cross-linguistic generalizations about word order.

3   Identifying constructions cross-linguistically

A variety of different problems arise in classifying languages according to various characteristics because of problems in deciding whether a construction in a given language should be considered an instance of a particular cross-linguistic type. Our discussion so far has assumed that we can identify instances of subjects, objects, genitives, postpositions, manner adverbs, standards of comparison, markers of comparison, and subordinators, but in practice there is considerable variation in what linguists count as instances of each of these, and one can face problems describing a language in deciding whether a particular construction in the language counts as an instance of the cross-linguistic category. In this section we discuss some of these problems and attempt to characterize, briefly, what are generally understood in the literature as instances of these categories, discussing some of the more frequent problems that arise.

3.1   Identifying the order of subject, object, and verb

3.1.1   Identifying subjects   Classifying a language as SV or VS seems to assume that the language has a clear instance of the category subject. There is an extensive literature discussing a variety of different possible problems with this assumption, and, for various problematic cases, different linguists have taken different positions as to what, if anything, should be counted as the subject. To some extent, one's decision on how to classify a particular language will depend on one's assumptions as to what is an instance of a subject.

Consider briefly the case of Tagalog. Schachter (1976, 1977, 1996), for example, argues that Tagalog lacks the category *subject*, that the properties that characterize subjects in other languages fail to isolate a single category in Tagalog. Let us suppose for the sake of argument that we accept Schachter's conclusion. How are we then to classify Tagalog in terms of the order of subject, object, and verb? Under the assumption that the category *subject* does not apply in Tagalog, there is a clear sense in which it would seem mistaken to classify the language as, for example, VOS (as is sometimes done). On the other hand, it is clearly the case that the single argument of intransitive verbs and the two arguments of

transitive verbs normally follow the verb in Tagalog, despite issues as to how to classify them. What this means is that there is a clear sense in which Tagalog is a verb-initial language, regardless of what if anything we call a subject.

3.1.2   The order of subject, object, and verb   The difficulty in classifying Tagalog as VOS or VSO reflects a more general problem in that there are many instances in which it is difficult to classify a language according to the six-way typology of SOV, SVO, VSO, VOS, OVS, and OSV. These problems arise, either because of difficulties deciding what if anything is subject or object, or, more commonly, because the flexibility of the language is sufficiently great that it is difficult to say that a single one of these orders is basic.

In many cases, however, such languages are more easily classifiable as SV or VS, or as OV or VO. One of the reasons for this is that transitive clauses containing a noun subject and a noun object do not occur very often in most languages, but clauses with just a noun subject or just a noun object are much more common. In many languages with flexible word order, frequency criteria will point to a classification of the language as, say SV and OV, in that subjects and objects more often precede the verb, but frequency criteria will leave the classification of the language as SOV more questionable. Descriptions of clause order in grammars of various languages often dwell too long on the problem of classifying the language as SOV or SVO, etc., and never even address questions of whether the language is OV or VO, even though the latter sort of question is often answered more easily. In addition, questions about the order of subject and verb in intransitive clauses are often ignored. But even if a language can be justifiably classified as SVO, it does not follow from this that the language is SV for intransitive clauses. Spanish is an example of a language which can be classified fairly uncontroversially as SVO, but whose classification as SV is more problematic due to the large number of situations in which VS order is employed in intransitive clauses. What this means is that one needs to ask, not only if the language is SV or VS, but whether there is a difference between transitive subjects and intransitive subjects in terms of their position with respect to the verb. (See Dryer (1997a) for further discussion of these issues.)

Identifying the order of subject, object, and verb involves identifying three different things: the order of subject and verb, the order of object and verb, and the order of subject and object. The arguments in the preceding paragraph argue that the first two of these are often easier to identify, while the third one is often more difficult to identify. If a language allows both orders of subject and object, answering this question is often fairly difficult. The question of identifying the basic order of subject, object, and verb in a language is often associated with the question of what role the order of subject and object plays in distinguishing which is subject and which is object in clauses containing a nonpronominal lexical subject and object. It should be noted, however, that in verb-final and

verb-initial languages in which pronominal arguments are expressed by verbal affixes, word order will not suffice for identifying the grammatical role of a single lexical noun phrase in a transitive clause.

3.1.3   Lexical noun phrases versus pronouns   The normal understanding of what we mean when we talk about the basic order of subject, object, and verb, or of just subject and verb, or of just object and verb, is the order when the subject or object is a noun, rather than a pronoun, or, more accurately, a lexical noun phrase, i.e. a noun phrase headed by a noun, rather than a noun phrase consisting of just a pronoun. In some languages, like English, pronouns exhibit a distribution that differs very little from that of lexical noun phrases, so that it makes little difference whether one includes pronouns or not in discussing the position of subjects and objects. But in many other languages, pronouns exhibit word order properties that differ considerably from lexical noun phrases, either because the syntactic rules of the language treat them differently, or because the pragmatic rules are such that their distribution is rather different. In Barasano, for example, a Tucanoan language spoken in Colombia (Jones and Jones (1991)), both preverbal and postverbal position are common for lexical subjects, but pronominal subjects normally follow the verb, as in (46).

(46)   yɨ-re   tudi-bɨsa-a-bã   ĩdã
         1SG-OBJ   scold-a.lot-PRES-3PL   3PL
         'they scolded me a lot'

The effect of this difference between lexical subject and pronominal subject is that, if one considers *all* subjects, both lexical and pronominal, then the most common order in Barasano is apparently OVS, but if one restricts attention to lexical subjects, the order is indeterminately SOV/OVS.

   It should be emphasized that, although the position of lexical subjects and objects is crucial in determining basic order at the clause level, the position of pronominal subjects and objects, if different from that of lexical subjects and objects, is just as important in giving a complete description of word order in the language.

3.2   *Identification of manner adverbs*

With many pairs of elements, a pervasive problem is that of to what extent one should use purely semantic criteria in identifying constructions and to what extent specific syntactic properties of the construction in the language are relevant. Many linguists have used purely semantic criteria in identifying constructions, but the semantic criteria employed are often strongly influenced by English translations and there is the danger of imposing English categories

on languages to which they do not apply. Consider, for example, the fact that verb-initial languages normally place manner adverbs after the verb. On the basis of purely semantic criteria, Jakaltek (Craig (1977)) would appear to be an exception to this, as in (47).

(47)   c'ul   xu   scañalwi naj
         good   did   dance   he
         'he danced well'

However, closer examination of the Jakaltek construction reveals that the word *c'ul* 'good' that translates into English as a manner adverb is not a modifier of the verb, but is actually itself the main verb in Jakaltek, while the verb that is the main verb in the English translation, *scañalwi* 'dance' is actually a subordinate verb. Classifying Jakaltek as AdvV would thus be very misleading: the word order in (47) actually reflects the fact that the main verb normally occurs first in Jakaltek and thus the order conforms to the general principles of word order in the language. In identifying something as a manner adverb in a language, there ought to be reason to believe that it is actually modifying the verb.

Otherwise, the identification of manner adverbs tends to be relatively unproblematic cross-linguistically. We have restricted discussion specifically to manner adverbs rather than other sorts of adverbs because in many languages other sorts of adverbs exhibit greater flexibility in their position with respect to the verb, and thus the correlation between the order of manner adverb and verb and the order of object and verb is stronger than it is with other adverbs. On the whole, however, other sorts of adverbs tend to exhibit a similar, though weaker, correlation. However, there are many languages in which temporal and locative adverbs defining the setting exhibit a tendency to occur in sentence-initial position, regardless of the order of object and verb.

Grammatical descriptions use the notion of manner adverb in different ways. Our assumption is that a manner adverb is an adverb modifying a verb denoting an event, and that the manner adverb denotes *how* the event took place. Prototypical manner adverbs are words corresponding to English *well*, *badly*, *quickly*, and *slowly*. On our use of the term, it does not apply to words like *immediately* (which is really a kind of temporal adverb) or *very* (which is a kind of degree word, discussed below in section 7.6).

3.3   Identification of prepositions and postpositions

A number of problems arise with identifying whether a language employs prepositions or postpositions. Again, the primary issue is that of to what extent semantic criteria are sufficient. Prototypical instances of adpositions are words that combine with noun phrases and that indicate the semantic relationship of

that noun phrase to the verb, as exemplified by the English word *with* in *He opened the door with a key.*

3.3.1   Adpositions versus case affixes   Perhaps the largest issue is whether the term ought to be applied to semantically similar morphemes which are affixes rather than separate words. For example, in Martuthunira, a Pama-Nyungan language spoken in western Australia (Dench (1995)), the meaning of the English preposition *toward* is expressed by a suffix on nouns, as in (48).

(48)   ngayu pamararri-lha ngurra-wurrini
        1SG   call.out-PAST   camp-toward
        'I called out towards the camp'

In the history of word order typology, such affixes have often been treated as adpositions; many of the languages classified as postpositional by Greenberg (1963) and Hawkins (1983) only have postpositions in the sense of having noun suffixes as in (48). One reason that the distinction between affixes like that in (48) and adpositional words is ignored by some is that such affixes often derive historically from separate adpositional words, and thus that, while the distinction may be valid synchronically, it is less important diachronically. A more common view, however, is that such morphemes should not be considered adpositions, since their position is defined in the morphology of the language in terms of their position with respect to the noun stem, rather than in the syntax in terms of their position with respect to the noun *phrase.*

3.3.2   Case affixes versus adpositional clitics   While it is probably best not to view case affixes as adpositions, it is also important to distinguish case affixes from adpositional clitics. Consider the object morpheme *-ga* in the example in (49), from Kanuri, a Nilo-Saharan language spoken in Nigeria (Hutchison (1976)).

(49)   kâm=ga   rúskəna
        man=OBJ 1SG.saw
        'I saw the man'

The morpheme *-ga* is attached to the noun in (49), and in such examples looks like a case suffix. However, when the noun is followed by a modifier, as in (50), it attaches to the modifier instead of the noun.

(50)   [kâm kúrà]=ga rúskəna
        man   big=OBJ   I.saw
        'I saw the big man'

The general rule, in fact, is that it attaches to whatever is the last word in the noun phrase. Its position is thus defined, not in the morphology, but in the syntax.

It is exactly like postpositional words, except that it attaches phonologically to the word that precedes it. For this reason, it is best viewed not as a case suffix, but as a type of postposition, namely a postpositional clitic. If the last element in the noun phrase happens itself to be a postposition, then we end up with a sequence of two postpositions, as in (51).

(51)   [fátò   [kâm kúrà]=ve]=ga rúskəna
           compound man big=GEN=OBJ I.saw
           'I saw the big man's compound'

The reason that two postpositions occur adjacent to each other in (51) is that the genitive postposition *-ve* is combining with the noun phrase *kâm kúrà* 'the big man' to indicate that it bears the genitive relation to the noun *fátò* 'compound', and the resultant noun phrase *fátò kâm kúràve* 'the big man's compound', is functioning as object of the clause and thus takes the object postposition *-ga.*

One can think of clitics like these postpositional clitics in Kanuri as morphemes which are syntactically separate words but phonologically like affixes in being attached to other words. They are syntactically separate words in that their position cannot be described in terms of the morphology of the language, but must refer to the syntax, in terms of their position relative to syntactic phrases. All evidence suggests that clitics behave the same as clear instances of words as far as word order correlations are concerned. Note that the English genitive clitic (*the Queen of England's crown*, *the woman I spoke to's hat*) counts as a postpositional clitic as well, and is quite parallel in various respects to the Kanuri object clitic.

Other examples of languages with postpositional clitics are illustrated in (52). The example in (52a) illustrates a locative postpositional clitic in Awtuw, a Sepik language of Papua New Guinea (Feldman (1986)), and the example in (52b) an ergative postpositional clitic in Thaayore, a Pama-Nyungan language of northeastern Australia (Hall (1972)).

(52)   a. [wutyæn dæni]=ke
          basket   INDEF=LOC
          'into a basket'

      b. [Pa:th nhaŋn]=man tha·th-iř   ru:ř   mant
          fire   his=ERG   burn-PUNCT insect small
          'his fire scorched the small grub'

In both cases, these markers are clitics rather than affixes, since they attach to whatever is the last word in the noun phrase.

Unfortunately, it is often unclear from many grammatical descriptions whether a morpheme that is called a case suffix is really a case suffix or a

postpositional clitic. Descriptions often refer to a morpheme as a case suffix and include it in the discussion of noun morphology, and it is only a brief mention elsewhere in the grammar, or sometimes only isolated examples, that reveal that it actually attaches to the last constituent of the noun phrase and thus is not a case suffix at all, but a postpositional clitic. This occasionally has ramifications for other aspects of the description of the language. For example, the fact that the clitic can appear on adjectives following the noun can lead some analysts to conclude from that that adjectives in the language are really nouns, when in fact no such conclusion is warranted. Note that applying this logic to the genitive clitic -*'s* in English would lead us to the bizarre conclusion that *singing* is a noun in examples like *the man that was singing's car* or that *to* is a noun in *the woman I spoke to's hat.* Unfortunately, a number of the papers in Plank (1995) apply the term *suffixaufnahme* (or 'double case') both to instances of multiple case affixes and to a number of instances of multiple postpositional clitics of the sort illustrated in the Kanuri example in (51) above. The two kinds of phenomena are really quite distinct, since the former is due to the nature of the morphology of the language, while the latter arises due to the coincidence of the syntax allowing two adpositions to occur adjacent to each other.

It is not always easy to distinguish adpositional clitics from case affixes. The most difficult instances are those in which noun phrases in the language are rigidly noun-final, in which all modifiers precede the noun, as in the Korean example in (53).

(53)   na-nun   ku   khun   kay-lul   po-ass-ta
       1SG-TOP that big   dog-ACC see-PAST-DECL
       'I saw that big dog'

The morpheme -*lul* 'accusative' in (53), like a variety of other 'case' morphemes in Korean, attaches to the noun, but since the noun is always the final element in the noun phrase, it is difficult to decide, on the basis of superficial evidence, between analysing the morpheme as a case suffix on nouns or as a postpositional clitic that always attaches to the last element of the noun phrase: in such a language, the last element of the noun phrase will always be the noun, so more sophisticated arguments may be necessary to choose between the two analyses. In the case of Korean, one piece of evidence that supports the postpositional clitic analysis is the fact that when the object involves a conjoined noun phrase, it is possible for the accusative morpheme to occur only once, at the end of the second noun phrase, as in (54).

(54)   na-nun   cakun kay-wa khun koyangi-lul po-ass-ta
       1SG-TOPIC small dog-and big   cat-ACC   see-PAST-DECL
       'I saw a small dog and a big cat'

3.3.3   Adpositions and relational nouns   The distinction between case affixes on the one hand, and postpositional words and postpositional clitics on the other, is only one of a number of possible problems identifying adpositions. A second common problem arises in many languages, in which some if not all of the words that translate as prepositions or postpositions are arguably really nouns. For example, England's (1983a) description of Mam (a Mayan language spoken in Guatemala) identifies a class of words she calls 'relational nouns', illustrated in (55).

(55)   ma   b'aj t-aq'na-7n   Cheep t-jaq' kjo7n   t-uuk'
        REC.PAST DIR 3SG.ERG-work-DIR José   3SG-in cornfield 3SG-with
           Xwaan t-e   xjaal
           Juan   3SG-for   person
        'José worked in the cornfield with Juan for the person'

The three Mam words corresponding to the English prepositions *in*, *with*, and *for* are all instances of what England calls relational nouns. She describes them this way since their morphology and the structure of the phrase consisting of the relational noun and the noun phrase that they combine with (e.g. *tuuk' Xwaan* 'with Juan') are identical to those of a noun phrase modified by a genitive, as in (56).

(56)   t-kamb'   meeb'a
        3SG-prize orphan
        'the orphan's prize'

If the words in question are really nouns, and if the construction is really an instance of a genitive construction, then classifying the language as prepositional may be an artefact of the English translation.

There are a number of other considerations, however, which make it less clear that it is a mistake to classify words like these relational nouns in Mam as prepositions. First, the fact that they are nouns does not entail that they are not prepositions. There might be language-internal criteria for distinguishing them as a subclass of nouns, in which case, we might say that there is a class of nominals in the language, two subclasses of which are prepositions and nouns. Even if the morphology and internal syntax of 'relational noun phrases' is like that of genitive noun phrases, there might be differences in their external syntax; it might be necessary, for example, to distinguish them from other noun phrases in describing the syntax of clauses.

A further consideration is that words that translate as prepositions in English often start out as nouns, but by processes of grammaticization gradually take on properties distinct from other nouns, even if they retain certain properties, such as nominal morphology, that reflect their historical origin. The general moral is that just because words in a language exhibit morphological properties

of nouns, it does not follow that they are not prepositions, for they may have acquired syntactic properties distinct from other nouns that reflect the relational functions associated with adpositions in other languages.

The processes of grammaticization whereby locational nouns become adpositions may eventually lead to loss of some nominal morphology. For example, in Kham (Watters (2002)), a Tibeto-Burman language spoken in Nepal, the normal genitive construction involves simultaneous genitive case marking on the genitive noun and pronominal possessive marking on the possessed noun, as in (57).

(57)   baza-e   o-kər
        bird-GEN 3SG.POSS-wing
        'a bird's wing'

There is a class of locational nouns which occur with both possessive prefixes and nominal case suffixes, as in (58).

(58)   a. nə-chĩ:-kə   b. o-ŋah-tə
           2SG-behind-at   3SG-front-on
           'behind you'   'ahead of it'

However, when they occur with a dependent noun, they do not occur with the possessive prefix, nor does the dependent noun occur in the genitive, as in (59).

(59)   a. hã: khɨ̃-kə   b. juhr   dũ:h-lə
           cliff foot-at   boulder   beneath-in
           'at the foot of the cliff'   'under the boulder'

Thus, despite their retaining some morphological characteristics as nouns, their lack of nominal morphology in some contexts reflects their partial grammaticization as postpositions. Synchronically, we can say that these words form a subclass of nouns, which we can call postpositions.

3.3.4   Languages without adpositions   While the majority of languages of the world appear to have adpositions, there are many languages that do not have words of this sort. In many Australian languages, this function is served by case affixes, as in the Martuthunira example cited above in (48). Such languages are not exceptions to the claim that OV languages tend to have postpositions. This claim is intended to be interpreted as saying that if a language is OV and if it has adpositions, then it will normally have postpositions rather than prepositions.

3.4   Identification of genitives

3.4.1   Alienable versus inalienable possession   There is relatively little cross-linguistic difficulty in identifying genitive constructions. One observation

that must be made, however, is that some languages employ distinct constructions for alienable and inalienable possession (see Dryer in vol. II, chapter 3), and, in a subset of these languages, the order of genitive and noun is different. For example, in Mallakmallak (Birk (1976)), a Daly language of north Australia, inalienable genitives employ GN order, as in (60a), while alienable genitives employ NG order, as in (60b).

(60)   a. yin^{y}a puntu   b. muyin^y yin^{y}a-nö
          man   head   dog   man-GEN
          'the man's head'   'the man's dog'

3.4.2   Lexical genitives versus possessive pronouns   Just as it is the case that discussions of the order of subject and verb are understood to refer to the order of lexical subject and verb, so too when we refer to the order of genitive and noun in a language, it is assumed that what is meant is a lexical genitive, one headed by a noun, rather than a pronominal genitive (also known as a possessive pronoun). In the majority of languages, the order is the same for lexical genitives and possessive pronouns (Poss), but in some languages their order is different. In French, for example, the order of lexical genitive and noun is NG, as expected of it as a VO language and as is illustrated in (61a), while the order of possessive pronoun and noun is PossN, as in (61b).

(61)   a. le   livre de Jean   b. son   livre
          the book   of Jean   3SG.POSS book
               N   G   Poss   N
          'Jean's book'   'his/her book'

Maranungku, a Daly language of north Australia (Tryon (1970b)), exhibits the opposite pattern, with the lexical genitive preceding the noun (at least for inalienable possession), as in (62a), and the possessive pronoun following the noun, as in (62b).

(62)   a. Micky piyamerr   b. piya ngany
          Micky hair   head my
          G   N   N   Poss
          'Micky's hair'   'my head'

4   Exceptions to word order generalizations

It must be stressed that the generalizations we discuss in this chapter are tendencies, and that there exist exceptions to most of them. For example, there are OV languages with prepositions, such as Kurdish (Abdulla and McCarus (1967)). The example in (63a) illustrates the OV order, while that in (63b) illustrates the use of prepositions.

(63) a. ʔəm pyaw-ə xənjə́r ʔə-froše
         this man-this dagger IMPERF-sell
              S   O   V
         'this man sells daggers'

      b. bïra-kə́-m   lə dïtróyt dàʔəniše
         brother-the-my in Detroit lives
                          Pr NP
         'my brother lives in Detroit'

And there are verb-initial languages with postpositions, such as Northern Tepe-huan, a Uto-Aztecan language spoken in Mexico (Bascom (1982)). The example in (64a) illustrates the VSO word order, and (64b) illustrates the use of postpositions.

(64) a. takávo   savîli [piidyúru ɨɨmádu ándɨrɨ̌ši] múi ɨ́ɨkoli . . .
         yesterday bought Peter   with   Andrew   many orange
                   V   S   O
         'yesterday, Peter and Andrew bought many oranges . . .'

      b. savîli ááni váík ímai giñ-ooñí-ga   vɨɨtári
         bought I   three squash my-wife-POSSD for
                                    NP   Po
         'I bought three squash for my wife'

There are also OV languages in which the genitive follows the noun, such as Arbore, a Cushitic language spoken in Ethiopia (Hayward (1984)). The example in (65a) illustrates the SOV word order of Arbore (the second word in (65a) is a nonverbal auxiliary particle varying for aspect and for the person and number of the subject) while (65b) illustrates the NG order.

(65) a. mo ʔíy   ḳor ḳúur-e   b. gaydan-ti géer
         man 3SG.DEF tree cut-PERF   hoe-POSSD old.man
         S   O   V   N   G
         'the man cut the tree'   'the old man's hoe'

Conversely, there are verb-initial languages with GN order, such as Yagua (D. L. Payne and Payne (1990)), as illustrated in (66).

(66) a. s-iimyiy Alchíco-níí   quiivạ   b. Tomáása rooriy
         3SG-eat   Alchíco-DEF.OBJ fish   Tom   house
         V   S   O   G   N
         'Alchico is eating the fish'   'Tom's house'

(The morpheme *níí* on the subject in (66a) is a pronominal clitic coreferencing the object.)

5   Other word order characteristics that correlate with the order of object and verb bidirectionally

In section 1, we saw five sets of elements whose order correlates with the order of verb and object. These sets of elements correlate *bidirectionally* in a sense that can be illustrated with adposition type. The correlation between OV order and postpositions is a strong tendency that can be stated by means of a bidirectional implicational universal: OV <=> Po, or 'A language is OV if and only if it is postpositional.' This is equivalent to saying 'If a language is OV, then it is postpositional, and if it is postpositional, then it is OV.' And by principles of logic, these also imply that if a language is VO, then it is prepositional, and if it is prepositional then it is VO. In this section we will examine a number of other pairs of elements whose order correlates bidirectionally with the order of object and verb. In section 6 below, we will examine some pairs of elements whose order correlates in a way I will characterize as unidirectional (rather than bidirectional), and in section 7, we will examine some pairs of elements whose order does not correlate with the order of verb and object at all.

5.1   *Verb and adpositional phrases*

The order of verb and adpositional phrase (PP) is usually the same as the order of verb and object. Thus the OV languages discussed above in section 1 generally place the adpositional phrase before the verb, as in Lezgian and Slave, as illustrated in (67).

(67)   a. [duxtur-r-in   patariw] fe-na
           doctor-PL-GEN to   go-AORIST
           PP   V
           'she went to doctors'

        b. Mary [Joe gha] ke   ehtsi̱
           Mary Joe for   slippers 3:is.making
                      PP   V
           'Mary is making slippers for Joe'

Conversely, VO languages, both verb-initial and SVO, normally place adpositional phrases *after* the verb, as in the examples in (68) from English and Fijian.

(68)   a. Mary cut the fish [with the knife]
                   V   PP

        b. au na   talai Elia [i 'Orovou]
           1SG FUT send Elia to 'Orovou
                      V   PP
           'I'll send Elia to 'Orovou'

5.2   *Verb and non-argument noun phrases*

Noun phrases that are not marked with an adposition but which are not syntactic arguments of the verb exhibit the same pattern in languages without adpositions in that they tend to occur on the same side of the verb as the object. For example, in Anguthimri, a Pama-Nyungan language of northeast Australia (Crowley (1981)), not only does the object normally precede the verb, but so do noun phrases that are not syntactic arguments. The example in (69) illustrates both the object and an instrumental NP preceding the verb. We use the symbol X to denote a non-argument NP, an NP that is not part of the lexical structure of the verb.

(69)   ʔwa-ra   bwaʔa   ba-gu   ṭa-na
       dog-ERG   meat   teeth-INSTR   bite-PAST
              O   X   V
       'the dog bit the meat with his teeth'

5.3   *Main verb and auxiliary verb*

In OV languages, auxiliary verbs normally follow the main verb, while in VO languages they normally precede. We saw above that Slave and Siroi are OV languages. The examples in (70) illustrate auxiliary verbs following the main verb in these languages.

(70)   a. bets'ę́   wohse   wolé   b. pasa   min-geŋ
          3.to   1SG.shout.OPT   be.OPT   talk   be-1PL.PAST
              V   Aux   V   Aux
          'I will shout to him/her'   'we were talking'

In contrast, English and Turkana are VO and AuxV, as in (71).

(71)   a. She   is   sleeping   b. kì-pon-i`   atɔ-mat-à
               Aux   V   1PL-go-ASP   1PL-drink-PL
                                Aux   V
                               'we shall drink'

The Turkana example in (71b) illustrates a verb *pon* 'go' functioning as a future auxiliary and preceding the main verb.

    The tendency for auxiliary verbs to occur on the opposite side of the verb from the object applies not only to auxiliary verbs coding tense or aspect but also to modal auxiliary verbs. The example from Lezgian (OV) in (72) illustrates a modal auxiliary expressing ability following the main verb.

(72) za-way   a   bejaburčiwal ex-iz   že-zwa-č-ir
     1SG-ADESS that shame   bear-INFIN can-IMPERF-NEG-PAST
                                    V   ModAux
     'I could not bear that shame'

The example in (73) from Moro, an SVO Kordofanian language spoken in Sudan (Black and Black (1971)), illustrates a modal auxiliary preceding the main verb.

(73) ña-gaməlu ña-gaber ña-gaṛəwaɖaṭa ña-gasa eɖe
     2PL-not.yet 2PL-not 2PL-able   2PL-eat meat.PL
                 NegAux   ModAux   V
     'you are still not able to eat meat'

The example in (73) from Moro also illustrates how, in languages in which negation is expressed by an auxiliary verb, such words exhibit the same tendency, following the verb in OV languages and preceding in VO languages.

   The expression 'auxiliary' is sometimes used to denote nonverbal particles which convey tense or aspect. The position of such particles does not correlate with the order of verb and object, as is discussed below in section 7.5.

5.4   Copula verb and predicate

In many but not all languages, clauses with nonverbal predicates require that a copula verb be used. The order of copula and predicate correlates with the order of verb and object, the copula generally following the predicate in OV languages, but preceding in VO languages. The order CopPred in a VO language is illustrated for English in (74a), while PredCop order in an OV language is illustrated in (74b) for Slave.

(74) a. Susan is   a doctor   b. ʔeyá hı̨lı̨
             Cop   Pred   sick   3SG.be
                                          Pred Cop
                                          'she is sick'

5.5   Question particles

Many languages distinguish polar (or 'yes/no') questions from corresponding declarative sentences solely by means of intonation. A few languages, like English, have syntactically different forms to signal such questions (e.g. *Is the dog barking?*). But many languages employ morphemes in polar questions to distinguish them from declarative sentences. While in some languages these are affixes, in other languages they are particles. There are a number of positions in which such question particles occur. In some languages, their position is

variable, depending on the focus of the question. In Turkish, for example, the
question particle immediately follows the word that is the focus of the question,
as illustrated in (75).

(75)   a. Sen kitap-lar-ı   al-dı-n   mı
           2SG book-PL-ACC take-PAST-2SG Q
           'did you TAKE the books?'

        b. Sen kitap-lar-ı   mı al-dı-n
           2SG book-PL-ACC Q   take-PAST-2SG
           'did you take THE BOOKS?'

In (75a), the focus of the question is the verb *aldın* 'take', while in (75b) it is
the object *kitapları* 'books'.

   There are many languages in which the question particle occurs in second
position, after the first constituent in the clause, as in !Xu, a Khoisan language
spoken in southern Africa (Snyman (1970)). In (76a), the question particle
occurs after the subject; but in (76b), in which there is an adverb preceding the
subject, the question particle follows this adverb and precedes the subject.

(76)   a. da'ama re ho n!eng
           child   Q see eland
           'does the child see the eland?'

        b. ǁe'ike re da'ama ho n!eng?
           today Q child   see eland
           'does the child see the eland today?'

   There are many other languages, however, in which the question particle
occurs either at the beginning of the sentence or at the end of the sentence, and
these two types correlate with the order of major clausal constituents. In OV
languages, these particles most often occur at the end of the sentence, as in the
example in (77) from Dolakha Newari, a Tibeto-Burman language spoken in
Nepal (Genetti (1994)).

(77)   Dolakhā khā tuŋ   lā-eu   rā
        Dolakha talk EMPH speak-3SG.FUT Q
        'will she speak the Dolakha language?'

In verb-initial languages, they most often occur at the beginning of the sentence,
as in the Lealao Chinantec example in (78).

(78)   sii^H ma^M-zá^L ka?^{M}ti^{LM} ku:^{H} kia:^{LH}a^H
        Q   PAST-run.out:3 completely money   POSS.1PL
        'has our money completely run out?'

While SVO languages pattern with verb-initial languages for most word order characteristics, they exhibit a pattern intermediate between OV languages and verb-initial languages with respect to question particles. Namely, SVO languages with initial question particles and SVO languages with final question particles are both common. The example in (79a), from Bagirmi, a Nilo-Saharan language spoken in Chad (Stevenson (1969)), illustrates an SVO language with a final question particle, while the example in (79b) is from Moro, illustrating an SVO language with an initial question particle.

(79)   a. i   ak   ŋʷon-ɨm   kau   lɛ?
            2SG see son-1SG   at:all Q
            'did you see my son at all?'

        b. an ña-gabəṭə nə-suk
            Q   2PL-go   to-Suk
            'are you going to the Suk?'

We use the expression 'question particle' here to denote particles in polar questions that are neutral with respect to what the answer might be. Many languages employ particles that occur in *leading questions*, in which the speaker makes an assumption as to what the answer will be, with a function analogous to the tag in English questions like *Mary is here, isn't she?* Such markers of leading questions appear to exhibit a tendency to occur at the end of sentences, regardless of the order of object and verb. For example, in Lealao Chinantec, illustrated above in (78) with a neutral polar question particle sii^Hat the beginning of the sentence, a question where a positive response is expected can be formed by means of the same particle at the *end* of the sentence accompanied by the negative word, as in (80).

(80)   na^M-ba^H ŋiú^H sïï^H ʔá:^{H}
        STAT-hit.3 house Q   not
                     S   Tag
        'the house was hit, was it not?'

It is also important not to confuse question particles with interrogative expressions in content questions, words corresponding to English words like *who* and *what.* The position of these is discussed below in section 8.1.

5.6   *Complementizer and clause*

Somewhat parallel to the case of adverbial subordinators is the order of complementizer and clause, where a complementizer is a word that signals the beginning or end of a complement clause, a clause functioning as object (or subject) of the verb in a higher clause. In English, for example, the

complementizer *that* occurs at the beginning of the clause, as in (81), illustrating the pattern that is typical for VO languages.

(81)   The teacher knows [that Billy ate the cookies]

This contrasts with the Slave example in (82), in which the complementizer *nį* occurs at the end of the clause, typical of OV languages.

(82)   [ʔelá táhła   nį̀]   kodeyihshá   yíle
         boat 3.land COMP   1SG.know   not
         Clause   Comp
         'I didn't know that the boat came in'

5.7   *Article and noun*

The order of article and noun exhibits a correlation with the order of verb and object, although the correlation is weaker than most of the other correlations discussed in this chapter. In particular, it is more common for the article to precede the noun in VO languages, as in English (*the dog*) and the Fijian example in (83a), but to follow the noun in OV languages, as illustrated in (83b) by the indefinite article in Kobon, an East New Guinea Highlands language.

(83)   a. 'eirau   'auta a   pua'a
           1EXCL.DU bring ART pig
                            Art   N
           'we (two) brought the pig'

        b. ñi   ap   wañib   i   ud   ar-nɨ m   Dusin laŋ
           boy INDEF string.bag   this take go-should.3SG Dusin above
           N   Art
           'a boy should take this string bag up to Dusin'
           Davies (1981)

In European languages, the term 'article' is used to denote words which code definiteness or indefiniteness and which, in some languages, vary with respect to other grammatical features of the noun phrase as well, such as case, gender, or number. Some languages elsewhere in the world employ words that do not vary for definiteness but which resemble articles in European languages in that they are words that are very common in noun phrases and which vary for grammatical features of the noun phrase (including number, case, gender), even if this does not include definiteness (see Dryer in vol. II, chapter 3, for further discussion). If we include such words in our understanding of the term 'article', i.e. if we treat definiteness not as a defining characteristic of articles,

but simply as a characteristic of articles in European languages, then words that introduce noun phrases in Cebuano, as in (84), count as articles.

(84) a. gi-palit   sa   babayi ang   saging
         OBJ.FOCUS-buy NONTOPIC woman TOPIC banana
                                [Art   N]   [Art   N]
         'the woman bought the bananas'

b. gi-sulat-an   ni   Maria si   Dudung ug   isturya
   LOC.FOCUS-write-LOC.FOCUS GEN Maria TOPIC Dodong INDEF story
                                      [Art N]   [Art   N]   [Art   N]
   'Maria wrote a story for Dodong'

Except for the indefinite object article *ug* in (84b), these articles in Cebuano do not vary for definiteness, but vary for a distinction between grammatical topic and various sorts of nontopics, and for common versus proper noun. (The nontopic actor Maria in (84) is marked with the same marker that marks genitive modifiers of nouns, hence the gloss 'GEN'.) Such articles appear to exhibit the same correlation with the order of verb and object that we have observed for articles coding definiteness or indefiniteness.

Many traditions categorize articles as a type of determiner, this class including as well such words as demonstrative modifiers of nouns (as in English *this book*). However, languages differ in whether articles and demonstratives belong to the same word class. In English, they do, appearing in the same determiner position at the beginning of noun phrases; in English, one cannot have both an article and a demonstrative (**the this book*). But in many other languages, articles and demonstratives are separate word classes. In Fijian, for example, the article precedes the noun while the demonstrative follows, as in (85).

(85) a   gone yai
      ART child this
      Art   N   Dem
      'this child'

In addition, unlike articles, demonstratives do not exhibit a correlation in their position with the order of object and verb, as discussed in section 7.2 below.

Some types of words, like demonstratives, are ones that are apparently found in all languages. Articles are a type of word where this is not so. While languages with articles are common, languages without articles are at least equally common, probably more so. But in addition to the weak correlation between the order of article and noun and the order of verb and object, there appears to be as well a weak correlation between the order of verb and object and *whether* the language employs articles. Namely, articles appear to be somewhat more

common in VO languages than they are in OV languages. A clear majority of OV languages appear not to have articles. Hence, when we say that OV languages tend to be NArt, what we really mean is that if a language is OV, and *if* it has articles, then it will tend to be NArt.

5.8   *Subordinate and main clause*

The position of adverbial subordinate clauses with respect to the main clause correlates with the order of object and verb, more often preceding the main clause in OV languages, and following in VO languages, although many languages exhibit considerable freedom in the position of subordinate clauses. English, for example, allows such clauses both before and after the main clause, as in (86), and it is not clear that one of these orders can be called basic.

(86)   a. Because it was raining,   the children came into the house
              Sub   Main

          b. The children came into the house   because it was raining
              Main   Sub

There is also some variation among different types of subordinate clauses. As Greenberg (1963) observed, conditional clauses exhibit a universal tendency to precede the main clause.

6   Word order characteristics that correlate with the order of object and verb unidirectionally

The pairs of elements discussed in the preceding sections are ones whose order correlates bidirectionally with the order of object and verb. What this means in effect is that, given the order of object and verb, one can predict that the language will probably have the other characteristics noted, and as well, given one of these other characteristics, one can predict the order of object and verb. For example, given the order OV, we predict VAux, and given VAux, we predict OV. The pairs of elements discussed in this section are not like this. We illustrate, with the first case we discuss, the order of relative clause and noun.

6.1   *Noun and relative clause*

Almost all VO languages place the relative clause after the noun, as illustrated in (87) for English, Fijian, and Tetelcingo Nahuatl.

(87)   a. the boy [that the dog bit]
                  N   Rel

b. a   pua'a ['eirau   'auta]
   ART pig   1EXCL.DU bring
        N   Rel
   'the pig which we (two) brought'

c. inu   ɔcintlɪ [tli   k-omwika-k]
   that water REL it-bring-PERF
        N   Rel
   'that water which he had brought'

However, among OV languages, both orders are about equally common. Examples of OV languages with RelN order are given in (88) from Lezgian and Awtuw. The Awtuw example in (88b) illustrates both the OV and the RelN order.

(88) a. [gada k'wal-iz   raq̃ur-aj]   ruš
         boy   house-DAT send-PTCPL girl
         Rel   N
         'the girl who sent the boy home'

      b. [rey   æye dək-ra-y-re]   rame-re wan d-uwp-o
         NONFEM.SG food ASP-eat-ASP-OBJ man-OBJ 1SG ASP-see-PAST
         Rel   N
         'I saw the man who is eating food'

The examples in (89) illustrate two OV languages with NRel order, Slave and Siroi.

(89) a. tthik'íhí [neyaa yet'ah golǫ   thehk'é sį́i]
         gun   2SG.son it.with moose 3.shot COMP
         N   Rel
         'the gun that your son shot the moose with'

      b. am   [ruga-nge ŋayong-ina] ta
         eye mud-SPEC ruin-3SG.PAST that
         N   Rel
         'the eye which the mud had injured'

We see therefore that three out of the four logical possibilities are common and that only one of the four is uncommon: OV&RelN, OV&NRel, VO&NRel are common, while VO&RelN is uncommon. We can describe this by means of a unidirectional implicational statement 'If VO, then NRel', or in its logically equivalent form 'If RelN, then OV'. What we *cannot* say is anything of the form 'If OV, then . . .' since, given that the order is OV, the two possibilities RelN and NRel are equally likely. Similarly, we cannot say anything of the form 'If

NRel, then . . .' since among NRel languages, many are OV and many are VO. In other words, the prediction goes in only one direction.

Contrast this with the situation involving the sort of bidirectional implicational generalizations that are possible, for example, with adposition type. In this case, only two of the four logical possibilities are common: OV&Po and VO&Pr. The other two possibilities are uncommon: OV&Pr and VO&Po. In this situation, the prediction goes in both directions: 'if OV, then Po' and 'if Po, then OV', as well as 'if VO, then Pr' and 'if Pr, then VO'.

In both types of situations, there is a correlation. It is clear that there is a correlation in the bidirectional case. In the unidirectional case, there is a correlation in the weaker sense that one order is significantly more common among OV languages than it is among VO languages. This contrasts with the cases we will look at in section 7 below in which there is no correlation at all, where all four types are common and the two orders are as common among OV languages as they are among VO languages.

The discussion above restricts attention to externally headed relative clauses, where the head is outside the relative clause and where it makes sense to talk about the order of the noun with respect to the relative clause. Slave employs both NRel externally headed relative clauses, as in (89a) above, but also internally headed relative clauses, as in (90).

(90)   [li̱ gah   hedéhfe i]   gháyeyidá
        dog rabbit chased   COMP 1SG.saw
        'I saw the dog that chased the rabbit'
        or 'I saw the rabbit that the dog chased'

Most languages with internally headed relative clauses are OV, as is Slave.

6.2   *Plural word and noun*

While the most common way to indicate plurality in a noun phrase is by means of an affix on the noun, a number of languages employ separate words to perform this function. Among VO languages with such plural words, both orders with respect to the noun are common, as illustrated in (91): (91a) illustrates PlurN order in Tahitian, an Austronesian language spoken on the island of Tahiti in the Pacific (Tryon (1970a)), and (91b) illustrates NPlur order in Tetun (Van Klinken (1999)), also an Austronesian language, but one spoken in Indonesia and East Timor.

(91)   a. te mau fare   b. hotu kakehe sia
           the PL   house   all fan   PL
               Plur N   N   Plur
           'the houses'   'all the fans'

Among the VO languages which place the plural word after the noun are some
where the plural word is a clitic which attaches to whatever is the last word in
the noun phrases, as in the examples in (92): (92a) is from Bagirmi and (92b)
from Margi (Hoffman (1963)), a Chadic language spoken in Nigeria.

(92)   a. [bɨs an ama]=ge   b. dàrà də̀zə̀=yàr
           dog of 1SG=PL   cap red=PL
           N   Plur   N   Plur
           'my dogs'   'red caps'

In contrast, all of the instances of OV languages with such plural words that
we are aware of place the plural word after the noun, as in the Siroi example in
(93).

(93)   kulim kat nuŋe
        sister PL his
        'his sisters'

We can summarize this distribution with the unidirectional implicational uni-
versal 'If a language is OV, then it will be NPlur.'

6.3   Intermediate unidirectional and bidirectional cases

6.3.1   Subordinator and clause   We have distinguished between two types
of correlations, bidirectional ones, where two of the four types are common
and the other two types less common, from unidirectional ones, in which three
of the four types are common and the fourth type less common. Because of
the vagueness of what it means to be common, there are in fact some cases
which might be classified either way. For example, we have treated the order
of adverbial subordinator and clause as a bidirectional correlation, since two of
the types, OV languages with final subordinators and VO languages with initial
subordinators, are more common than the other two possibilities. However,
of the two other possibilities, one is much rarer than the other. Namely, OV
languages with initial subordinators are much more common than VO languages
with final subordinators.

An example of an OV language in which subordinators occur at the beginning
of the clause is Latin, as in (94).

(94)   ubi   [puella-m audī-v-ī]
        when   girl-ACC hear-PERF-1SG
        Subord   Clause
        'when I heard the girl'

An example of a VO language with clause-final subordinators is Buduma, a Chadic language spoken in Sudan (Lukas (1939)). The SVO order of Buduma is illustrated in (95a), the clause-final subordinator in (95b).

(95) a. kugúi̯ a-tái̯   ámbai̯
          hen   3SG.MASC:PRES-lay egg
          S   V   O
          'the hen lays eggs'

      b. [dōmo hámera ná-ci-n]   ga
          1SG   cold   3SG.MASC:PAST-grip-1SG.OBJ since
          Clause   Subord
          'since I am cold'

OV languages like Latin are not uncommon: the implicational universal 'If OV, then final subordinator' is true for approximately 75 per cent of OV languages. In contrast the implication 'If VO, then initial subordinator' is apparently true for over 95 per cent of VO languages. Treating a case like this as a bidirectional correlation obscures the fact that one of the two less frequent types is much more common than the other, while treating it as a unidirectional correlation obscures the fact that two of the types are more common than the other two types.

6.3.2   Complementizer and clause   The order of complementizer and clause is similarly an intermediate case: of the two less frequent types, OV&Comp-Clause is not uncommon (found in over 20 per cent of OV languages), while we are aware of no instances of VO&ClauseComp languages. The example in (96), from Harar Oromo, a Cushitic language spoken in Ethiopia (Owens (1985)), illustrates an instance of an OV language with an initial complementizer.

(96) [akká-n   d'ufé-n]   beexa
      COMP-1SG came-1SG   know
      Comp   Clause
      'I know that I came'

Harar Oromo is somewhat atypical among OV&CompClause languages in that the complement clause occurs in normal object position before the verb, as in (96). More commonly in such languages, complement clauses follow the verb, contrary to the normal OV word order, as in the example in (97) from Hindi.

(97) aurat ne   kahaa [ki   aadmii ne   patthar maaraa]
      woman ERG said   COMP   man   ERG rock   threw
                         Comp   Clause
      'the woman said that the man threw the rock'

In both of these cases, the implicational relationship between the order of object and verb and the order of complementizer and clause is bidirectional, but it is much stronger in one direction than in the other, meaning that there is an asymmetry that resembles the unidirectional implications.

7   Word order characteristics that do not correlate with the order of object and verb

There are a number of word order characteristics that do not correlate cross-linguistically with the order of verb and object, where both orders are common in both OV and VO languages, or at least where there is no difference between OV and VO languages with respect to the frequency of the two orders of these other pairs of elements. The existence of such word order characteristics has often been overlooked in the literature. We discuss six such pairs of elements in this section.

7.1   Adjective and noun

7.1.1   The absence of a correlation with the order of object and verb   It is often mistakenly thought that the order of adjective and noun correlates with the order of object and verb, but it is now known that this is not the case (see Dryer (1988, 1992)). It is often thought that OV languages tend to be AdjN and that VO languages tend to be NAdj, but it turns out in fact that this is not so, that NAdj is somewhat more common than AdjN among both OV and VO languages. Part of the source of this problem is that the languages in the sample used by Greenberg (1963) suggested that verb-initial languages tend to be NAdj, but in fact this turns out to be an accidental property of the six verb-initial languages in his sample, and AdjN order is as common in verb-initial languages as it is in SVO and OV languages. Another source of the mistaken impression many linguists had about AdjN order in OV languages is that among the OV languages of Europe and Asia, AdjN order is much more common than NAdj order. This turns out, however, to be an idiosyncracy of Eurasia: outside of Eurasia, NAdj is clearly more common than AdjN among OV languages. The examples in (98) illustrate OV&AdjN and OV&NAdj order in Lezgian and Slave respectively.

(98)   a. i   güzel   cükw-er   b. tli̧   nechá
              this beautiful   flower-PL   dog big
                     Adj   N   N   Adj
              'these beautiful flowers'   'big dog'

Turning to verb-initial languages, the examples in (99) illustrate NAdj order in Fijian and Lealao Chinantec.

(99)   a. a   'olii loa   b. mï^VH-kuï:^M tia:^M
            ART   dog   black   CLSFR-corn   white
                   N   Adj   N   Adj
            'black dog'   'white corn'

Rukai, spoken in Taiwan (P. Li (1973)), is like Fijian in being a verb-initial
Austronesian language, but differs in being AdjN; (100a) illustrates the verb-
initial order, while (100b) illustrates the AdjN order.

(100)   a. wauŋul sa   acilay kay   maruḍaŋ
            drank   INDEF.ACC   water   this.NOM   old.man
            V   O   S
            'this old man drank water'

        b. kayvay maḍaw daan
            this   big   house
                     Adj   N
            'this big house'

Mezquital Otomi, an Oto-Manguean language spoken in Mexico (Hess (1968)),
is a second example of a verb-initial language with AdjN order; (101a) illustrates
the verb-initial order, while (101b) illustrates the AdjN order.

(101)   a. pě?ca ?na ra   ngǔ   nú?ạ ra   rǐko
            has   one ART house that ART rich.man
            V   O   S
            'that rich man has a house'

        b. ra   zí   zụ?wẹ́
            ART little   animal
                  Adj   N
            'the little animal'

7.1.2   Identifying adjectives   In characterizing the order of noun and adjec-
tive in a language, it is important to understand that what is at issue is the order
of a noun and an adjective that is *modifying* the noun, in an attributive function
within the same noun phrase, and not the order of a noun (phrase) functioning as
subject and an adjective functioning as predicate. Thus (99a) above illustrates
the NAdj order of Fijian, while (102) does not illustrate the order AdjN, but
rather the fact that in a clause in which the adjective is predicate, the predicate
precedes the subject.

(102)   e   loa.loa a   'olii yai
        3SG big   ART dog   this
              Pred   S
        'this dog is big'

Characterizing the order of noun and adjective in a language would seem to assume that one can identify a class of words that can be described as adjectives. There are a number of problems with this that arise in different languages. First, in the broadest sense of the word, adjectives include demonstrative 'adjectives' and numerals. However, the term 'adjective' is usually understood to denote what are sometimes called 'descriptive adjectives', words modifying nouns that denote properties of the referent of the noun phrase, the prototypical properties being ones with meanings like 'big', 'small', 'good', 'bad', 'old', 'new', and colours (see Dixon (1977b)). In addition to demonstratives and numerals, this excludes meanings like 'other', 'same', and 'such', which in some languages exhibit different word order properties from descriptive adjectives.

Another problem that arises in identifying adjectives is that in many languages the meanings in question are expressed by words that belong either to the class of verbs in the language concerned or to the class of nouns (see Dryer in vol. II, chapter 3, for further discussion). We thus encounter again the question of to what extent the categories assumed in word order typology are semantic and to what extent they are motivated as categories within each language. We follow here the general practice in word order typology of assuming a semantic notion of adjective, so that we include words that in some languages belong to the class of nouns, in others to the class of verbs

In some languages, although adjectives are a subclass of verbs, they may exhibit differences from other verbs in terms of their position relative to the noun. In Hanis Coos, for example, adjectival verbs modifying a noun normally precede the noun, as in (103a), while other verbs modifying a noun normally follow the noun, as in (103b).

(103)   a. tsä´yuxᵘ   tcîcī´mîł
            small   spruce.tree
            V   N
            'a small spruce tree'

        b. tE   to´qmas   k!a´wat
            the woodpecker peck
                 N   V
            'the woodpecker who was pecking at it'

What this means is that it does not follow from the fact that adjectives are verbs in a language that their position relative to the noun is necessarily governed by the same principles as that of other verbs. In fact, the positional properties of the words with adjectival meaning could be the basis for saying that they are a distinct word class, and calling them adjectives.

Finally, it should be noted that there are languages which can superficially be characterized as NAdj or AdjN, but in which such a characterization is highly misleading because the relation of the noun and adjective is not one in which

the adjective is modifying the noun, in attributive function. For example, (104a) from Kutenai appears to illustrate AdjN order, with the adjective *kwiɬqa* 'big' preceding the noun *t'awu* 'gun'. However, the structure of the noun phrase in (104a) is actually that of an internally headed relative clause, in which the adjective is a verb functioning as the predicate in the relative clause and the noun is the subject of that predicate. The apparent AdjN order in (104a) thus reflects the more general fact that the normal order of clauses in Kutenai is VS, as in (104b).

(104)   a. k-wiɬqa   t'awu   b. qa·nax-i   skinku¢
            SUBORD-big gun   go-INDIC   coyote
            V   S   V   S
            'the big gun'   'Coyote went along'

In other words, the noun in (104a) is not the head, with the adjective as modifier, but the so-called "adjective" is the head (assuming the verb is head of the clause) and it is the noun which is a dependent, more specifically the subject. If we restrict classification of the order of noun and adjective to cases in which the adjective is modifying the noun, then cases like Kutenai should be excluded. It is possible that there are other languages which have been described as AdjN or NAdj in which the structures in question are really internally headed relative clauses.

7.2   *Demonstrative and noun*

Demonstrative modifiers of nouns, like adjectives, are common either before the noun or after the noun among both OV and VO languages, though in both types of languages DemN order is slightly more common. The example in (105) illustrates DemN order for Lezgian.

(105)   a   insan-ar
            that   human-PL
            Dem N
            'those people'

The examples in (106) are from Canela Krahô, a Je language spoken in Brazil; (106a) illustrates the OV order, while (106b) illustrates the NDem order.

(106)   a. wa ha pĩxô jũhkà   b. rop ita
            1SG FUT fruit buy   dog this
            S   O   V   N   Dem
            'I will buy fruit'   'this dog'

The examples in (107) illustrate two verb-initial Oceanic languages with DemN and NDem order respectively, namely Tahitian and Fijian.

(107)   a. 'Ua   ti'i   'outou 'i   tēra   tiare.
            COMP   pick   2PL   PREP   that   flower
                                            Dem   N
            'you (pl.) picked that flower'

        b. a   gone   yai
            ART   child   this
                    N   Dem
            'this child'

7.3   Numeral and noun

Both NumN and NNum order are common among OV and VO languages, the
examples in (108) illustrating this for two OV languages, Lezgian and Slave,
the examples in (109) for two verb-initial languages, Lealao Chinantec and
Turkana.

(108)   a. i   wad   cük   b. dene   nákee
            this   five   flower   person   two
                    Num   N   N   Num
            'these five flowers'   'two people'

(109)   a. tu̠:^{L} ʔí^{VH} b. ŋa-kine-i`    ŋa-arey`
            two   place   PL-goat-PL   PL-two
            Num   N   N   Num
            'two places'   'two goats'

Note that in describing the order of numeral and noun, it is the order of
cardinal numeral (e.g. English *two*, *three*) and noun that is intended, rather than
the order of *ordinal* numeral (e.g. *second*, *third*) and noun. In some languages,
the position of cardinal numeral and ordinal numeral are different. For example,
in Gude, a Chadic language spoken in Nigeria (Hoskison (1983)), the cardinal
numeral follows the noun, as in (110a), while the ordinal numeral precedes the
noun, as in (110b).

(110)   a. mbusə   pu'   b. tufə-nə   nga   tihinə
            pumpkin ten   five-ORD   of   horse
            N   Num   Ord   N
            'ten pumpkins'   'fifth horse'

7.4   Negative particle and verb

We restrict attention here to negative morphemes that are neither affixes on
verbs, nor negative auxiliaries, discussed above in section 5.3. While both orders

of negative particle and verb are found in both OV and VO languages, preverbal position is more common. The examples in (111) from Slave and Waskia, a Trans-New Guinea language (Ross and Paol (1978)), illustrate the two orders of negative particle and verb in OV languages.

(111)   a. dene   gháyeyídá yíle
           person 1PL.see   not
                          V   Neg
           'we didn't see anyone'

        b. ane yu   me   nala bage-sam
           1SG water not   drink stay-PRES.1SG
                         Neg V
           'I never drink water'

The examples in (112) illustrate the two orders in two SVO languages, Bagirmi and Tetelcingo Nahuatl.

(112)   a. deb-ge   tol tobio li
           person-PL kill lion   not
                      V   Neg
           'the people did not kill the lion'

        b. amo nɪ-k-matɪ koš   ok   om-pa-ka . . .
           not   I-it-know whether still MED-at-be
           Neg V
           'I don't know if he's still there . . .'

Note that negative particles preceding the verb most often occur immediately before the verb, while negative particles following the verb in SVO languages often occur in clause-final position, as in (112a).

    Although the order of negative particle and verb does not correlate with the order of object and verb, it does actually correlate weakly with the order of subject and verb, in that the preverbal preference is particularly strong in verb-initial languages, and there are very few known verb-initial languages with postverbal negative particles. The examples in (113) from Lealao Chinantec and Yagua (D. L. Payne (1990)) illustrate two verb-initial languages with preverbal negative particles.

(113)   a. ʔa^{L}ʔe^M maʔ^L-líʔ^{L}i   b. néé ra-vyạạta buyạạ
           not   ASP-remember.1SG   not 1SG-want manioc.beer
           Neg V   Neg V
           'I no longer remember'   'I don't want manioc beer'

7.5 *Tense–aspect particle and verb*

By tense–aspect particles we mean uninflected words that indicate tense or aspect, similar to auxiliary verbs, but nonverbal. Such words are sometimes referred to as auxiliaries, particularly in languages in which they are clitics, or clitic clusters, that occur in second position in the clause. In languages in which their position is defined relative to the verb (as opposed to being in second position), they tend to precede the verb in both OV and VO languages. They do follow the verb slightly more often in OV languages, but this difference is sufficiently weak that we treat them here as not correlating with the order of object and verb. The examples in (114) illustrate the two orders in two OV languages spoken near the mouth of the Amazon River in Brazil, Urubu-Kaapor (Kakumasu (1986)) and Canela Krahô.

(114) a. kase   a-'u   ta   b. wa ha   pĩxô jũhkà
           coffee 1SG-drink FUT   1SG FUT fruit buy
                   V   T/A   T/A   V
           'I will drink coffee'   'I will buy fruit'

Note that the future particle in the Canela Krahô example in (114b) does not occur adjacent to the verb, but immediately after the subject. The examples in (115) illustrate the two orders in two VO languages, preverbal position in Mam and postverbal position in Bagirmi.

(115) a. ma   kub' ky-tzyu7n   xiinaq cheej
           REC.PAST DIR   3PL.ERG-grab man   horse
           T/A   V
           'the men grabbed the horse'

       b. bɨs sa   ja   ga
           dog eat meat COMPLETIVE
                V   T/A
           'the dog has eaten the meat'

Note that the postverbal completive particle in the Bagirmi example in (115b) occurs not only after the verb, but after the object, at the end of the clause.

7.6 *Degree word and adjective*

A final pair of elements whose order does not correlate with the order of verb and object is what are variously called degree words, intensifiers, or adverbs, words modifying adjectives that are analogous in meaning to English words like *very*, *more*, *rather*, *somewhat*, and *slightly*. Again, this is another point on which Greenberg's thirty-language sample was misleading, since the verb-initial languages in his sample were primarily AdjDeg. In fact, both orders

are common among verb-initial languages, and, if anything, the order DegAdj is slightly more common. The examples in (116) illustrate two verb-initial languages of each sort, Lealao Chinantec with DegAdj order, and Ocotepec Mixtec (Alexander (1988)) with AdjDeg order.

(116) a. díʔ^{VH} li:ʔ^H b. káhnú ndāsi
           very   pretty   big   very
           Deg   Adj   Adj   Deg
           'very pretty'   'very big'

Both orders are also common among SVO and OV languages.

   It should be mentioned that, in many languages, degree words do not behave as a grammatically well-defined class, and often some degree words precede the adjective while others follow, within the same language. The examples in (117) from Chrau, a Mon-Khmer language spoken in Vietnam (Thomas (1971)), illustrate two different degree words with different positions relative to the adjective.

(117) a. mo'yăh maq   b. maq trôq
           very   big   big   extremely
           Deg   Adj   Adj   Deg
           'very big'   'extremely big'

8   Other typological characteristics correlating with the order of object and verb

The correlations with the order of object and verb discussed in sections 1, 5 and 6 above all involve pairs of elements where one order is more common than the other order in OV languages as compared with VO languages. In this section, we discuss a number of typological characteristics that do not involve the order of two elements but that do appear to correlate with the order of object and verb. First, however, we should mention two such characteristics that we discussed briefly above. One of these is that internally headed relative clauses rarely occur outside of OV languages. The other is that articles are apparently used more often in VO languages than they are in OV languages.

8.1   *Position of interrogative expressions in content questions*

In section 5.5 above, we discussed the position of question particles, words signalling polar questions. These need to be distinguished from interrogative words or expressions that occur in content (or 'wh') questions. The examples in (118) from Otomi illustrate a question particle in a polar question and an interrogative expression in a content question, respectively.

(118)   a. ha   mǎ   gixɔ̌   nú   ra boxą́
           Q   going.to you.will.take.it that the money
           'Are you going to take the money?'

        b. té   gíhoní   wa
           what you.seek.it here
           'What are you looking for here?'

We saw in section 5.5 that question particles can occur in various positions, with initial position correlating with verb-initial languages and final position correlating with OV languages. The position of interrogative expressions in content questions also correlates with the order of object and verb: in verb-initial languages, such expressions most commonly occur at the beginning of sentences (and thus the verb is not initial in such sentences), while in OV languages, they tend most often to occur *in situ*, the same position in which a corresponding noninterrogative expression would occur. The example in (118b) from Otomi illustrates this initial position in a verb-initial language. The example in (119) from Slave illustrates *in situ* position in an SOV language: the interrogative expression is functioning as object and thus occurs after the subject and before the verb, where objects normally occur in Slave.

(119)   David ʔayíi ehtsi
        David what 3.make
        'what did David make?'

It should be stressed that this correlation is not as strong as some of those we have discussed and there are many exceptions. Fijian, which we have used to exemplify many characteristics typical of verb-initial languages, is exceptional in this respect. In (120), the interrogative occurs in subject position, in VOS order.

(120)   e   sabici i'o   o   cei
        3SG hit   2SG.OBJ ART who
        'who hit you?'

Unlike most of the word order characteristics we have discussed, in which SVO languages pattern like verb-initial languages, both types of content questions are common among SVO languages. English is an example of an SVO language in which the interrogative expression occurs at the beginning of the sentence. Hmong Njua is an example of an SVO language in which the interrogative expression occurs *in situ*, as in (121).

(121)   nwg moog ghov twg
        3SG go   where
        'where is he going to?'

Note that the interrogative expression is often an entire phrase, and in languages in which the interrogative expression occurs in initial position, the interrogative word may occur later in the phrase, so that the interrogative word itself is not initial, although the phrase is. For example, in (122) from Songhai, a Nilo-Saharan language spoken in west Africa (Prost (1956)), the interrogative phrase in initial position is *koyra foyan ga* 'in which villages', literally 'villages which in', in which the interrogative word *foyan* 'which' follows the noun (the general position for most noun modifiers in Songhai) and the entire postpositional phrase occurs in initial position.

(122)   koyra   fo-yan   ga   n   ga   bisa
         village which-PL in   2SG INCOMP pass
         'by which villages did you pass?'

8.2   *Affix position*

Although there are many differences among affixes of different sorts, there is overall a tendency for suffixes to be associated with OV languages, prefixes with VO languages. This is a unidirectional correlation, however, in that three of the four types are common, suffixes in OV languages, suffixes in VO languages, and prefixes in VO languages. In other words, we can say that OV languages more commonly have suffixes, but we cannot say that VO languages more commonly have prefixes. However, if a language is exclusively suffixing, if all affixes in the language are suffixes, the language is more likely to be OV. This correlation is not a strong one, and prefixes in OV languages are not at all rare.

8.3   *The use of case in distinguishing transitive arguments*

Languages employ a variety of means for distinguishing the two arguments in a transitive clause. One means is to mark one or both of them with a case affix or adposition. There appears to be a weak correlation by which OV languages employ such case markers more often than VO languages, more specifically one can say that they are most common in OV languages, next most common in verb-initial languages, and least common in SVO languages.

9   **Other sorts of implicational generalizations**

The best-known generalizations in word order typology have been ones relating the order of certain pairs of elements to the order of object and verb. With bidirectional correlations, this also means that pairs of elements whose order correlates with the order of verb and object correlate with each other. For example, languages which are postpositional tend to be GN and vice versa. But there

are also correlations between pairs of elements neither of which, or only one of which, correlates with the order of object and verb. For example, when we consider the order of adjective and noun and the order of demonstrative and noun, there are four logically possible combinations: DemN&AdjN, DemN&NAdj, NDem&NAdj, and NDem&AdjN. However, of these four types, the first three are common while the last one is uncommon. This can be described in terms of the unidirectional implicational universal 'If NDem, then NAdj' (or equivalently 'If AdjN, then DemN').

Greenberg (1963) and Hawkins (1983) discuss other possible universals that refer to three or more elements. For example, Greenberg's Universal 5 states 'If a language has dominant order SOV and the genitive follows the noun, then the adjective likewise follows the noun.' Note that the set of languages defined by the antecedent clause here is already somewhat small, since the genitive normally precedes the noun in SOV languages. Most of the universals of this form that have been proposed do appear to have some exceptions. Tigre (Raz (1983)), a Semitic language spoken in Eritrea, is an apparent exception to Greenberg's Universal 5. Examples illustrating these properties are given in (123): (123a) illustrates the SOV word order; (123b) illustrates the NG order; and (123c) illustrates the AdjN order.

(123) a. . . . ḥatte ʔəssit   walat-ʔəsrael ḥəṣān waldat
              one   woman Israelite   boy   begot
                    S   O   V
       '. . . an Israelite woman begot a boy'

       b. walat   farʕon
          daughter   Pharaoh
          N   G
          'the daughter of the Pharaoh'

       c. la-gəndāb   ʔənās
          the-old   man
          Adj   N
          'the old man'

10   Order among elements at the same level

We have discussed above the order of the noun relative to various modifiers, but questions also arise about the order among modifiers. And while noun phrases containing three or more modifiers are likely to be rather unnatural (though languages appear to differ in how tolerant they are of multiple modifiers), questions about the order of single pairs of elements can usually be answered. In other words, it may be unnatural in a language to express a noun phrase with three modifiers, as in English *these three brown books*, but the order

Dem-Num-Adj-N can be determined by examining one pair at a time (*these three books*, *these brown books*, *three brown books*). And as with other pairs of elements, languages vary as to whether the order is rigid, and, if it is flexible, whether one order is preferred or not, and what might determine the order. While the order is quite rigid in English, for example, it is more flexible in Japanese.

One cross-linguistic generalization governing the order of Dem, Num, and Adj is that when all three appear on the same side of a noun and one order is preferred, the demonstrative typically is furthest from the noun and the adjective closest, with either Dem-Num-Adj-N or N-Adj-Num-Dem order. Compare the English example in (124a) with the example in (124b) from Ambai (Silzer (1983)), an Austronesian language spoken in Irian Jaya in Indonesia.

(124) a. these three   brown books
          Dem Num   Adj   N

       b. dian katui siri   nani
          fish small one   that
          N   Adj   Num Dem
          'that one small fish'

The same generalization applies to any pair of these elements if two occur on one side of the noun. Thus if a language places both the demonstrative and the numeral before the noun and the adjective after the noun and if there is a preferred order for the Dem and Num, that order will typically be Dem before Num. Greenberg (1963) noted the existence of some languages which violate this in which all three of these elements follow the noun, but in the opposite of the expected order, namely N-Dem-Num-Adj, but such languages do not appear to be much more common (if at all) than other types of exceptions. The example in (125) from Moro illustrates a language of this sort.

(125) maj-anda ildi   iɽəjin   l-amənu   l-oɽra
       man-PL   this:NC$_{10}$.PL three:NC$_{10}$.PL NC$_{10}$.PL-black NC$_{10}$.PL-big
       N   Dem   Num   Adj   Adj
       'these three big black men'

An example of another type of exception is Nkore-Kiga, a Bantu language spoken in Uganda (Taylor (1985)), in which the order is N-Poss-Dem-Adj-Num, as illustrated in (126).

(126) ekitabo kyawe ekyo ekihango ekimwe . . .
       book   your   that large   one
       N   Poss   Dem Adj   Num
       'that one large book of yours. . . .'

Questions about the order of elements at the same level also apply at the clause or verb phrase level. Thus, identifying a language as OV and XV, where X stands for any adpositional phrase or noun phrase that is not an argument of the verb, leaves open the question of the possible order of the object with respect to other elements before the verb. In some languages, their order is flexible, in others the preferred order is XO, while in others the preferred order is OX. For example, the preferred order in Sanuma, a Yanomami language of Brazil (Borgman (1990)), is SXOV, as in (127a), while the preferred order in West Greenlandic (Fortescue (1984)) is SOXV, as in (127b).

(127)   a. pata   töpö-no sokopi a-nö   wale   kökö sepalöma
               old   3PL-ERG lance   3SG-INSTR peccary 3DU   killed
               S   X   O   V
             'the old people killed the peccary with lances'

    b. imaallaat filmi taanna Nuum-mi taku-ara
           luckily   film   that   Nuuk-LOC see-1SG.3SG.INDIC
                     O   X   V
             'luckily, I saw that film in Nuuk'

A similar three-way typology of flexible, OX, and XO, applies to VO languages, except, here, languages of the type OX (and thus VOX) are overwhelmingly the most common, and no language with basic order VXO is known to us.

11   Languages with flexible word order

In section 2, we discussed the problem of identifying a basic order for elements in languages in which more than one order is possible. It is sometimes mistakenly thought that word order typology is not relevant to languages with flexible word order. We have discussed one reason why this view is mistaken, namely that often, despite the flexibility, arguments can be given for treating one order as basic by one or more of the criteria. But there are additional reasons why word order typology is relevant to such languages. First, in many languages in which word order is flexible for some elements, it is less flexible for others. For example, in Tiwi, a language of northern Australia (Osborne (1974)), the order of elements at the clause level is quite flexible, but within the noun phrase, the order of modifiers with respect to the noun is fairly rigid.

Second, languages with highly flexible word order are themselves a linguistic type. There are many questions, largely still unanswered, about what generalizations can be made about such languages. It appears to be the case that word order flexibility is more common at the clause level than at the phrase level, so that we can say that if a language has flexible word order at the phrase level, then it will have flexible order at the clause level. There is also some reason

to believe that there is some correlation between polysynthesis and word order flexibility, but the exact nature of this correlation remains to be investigated. And it may be the case that languages with highly flexible word order tend more often to exhibit word order characteristics associated with OV languages rather than those associated with VO languages.

Third, a largely unexplored area of word order typology is what subtypes may exist among languages with flexible word order. To what extent is word order in such languages determined by pragmatic principles? And what factors other than pragmatic principles determine word order in such languages? And in so far as word order is determined by pragmatic principles, to what extent do the pragmatic principles vary among such languages, to what extent can we identify a further typology of the ways in which pragmatic principles determine word order?

Perhaps the most important observation to be made is that in describing a language with flexible word order, one should identify minimally just where the word order is flexible and where it is not, if possible what orders are more common, and ideally what factors govern the choice between alternative word orders. The latter task is usually very difficult, and there is considerable terminological confusion and vagueness in the literature discussing notions that may be relevant in different languages.

12   Typological versus language-particular description of word order

The various word order characteristics discussed in this chapter provide a basis for minimally characterizing word order in a language, but there is usually much more to be said about word order in a language beyond simply identifying a language with respect to these characteristics. For one thing, most languages allow both orders for some pairs of elements, and often the full description of the various factors relevant will be quite complex. The following examples from English illustrate just some of the complexity surrounding the position of adjectives and adjective phrases relative to the noun: *the interesting man*, **the man interesting*, *the only interesting man*, *the only man interesting*, **the afraid man*, **the man afraid*, *the only man afraid*, *a man more interesting than the mayor*, **a more interesting than the mayor man*, *a more interesting man than the mayor*, etc.

Second, there are many more fine-tuned questions that can be asked with respect to certain categories of words that are found widely among languages in the world that do not always fall cleanly into the categories we have discussed. At the level of noun modification, for example, one can ask about the position of interrogative modifiers with meanings such as 'which', 'what sort of', 'how many', and 'whose'. In most languages, such elements occur in

the same place in the noun phrase as corresponding noninterrogative words, but there are exceptions. For example in Ocotepec Mixtec, most modifiers, including demonstratives and adjectives, follow the noun, as in (128a), but the interrogative modifier meaning 'which' or 'what' precedes the noun, as in (128b).

(128)   a. ñūu   lūlí   ñúkwán   b. na   teē
           town   little   this   what   man
           N   Adj   Dem   Int   N
           'this little town'   'what man?'

And third, most languages have some words whose position is different from that of other words in the language and which simply require special description. For example, Mam (verb-initial) employs directional particles which precede the verb, like the particle *jaw* 'up' in (129); its behaviour is unlike that of adverbs, for example, which follow the verb in Mam.

(129)   ma   jaw b'iit'j
           REC.PAST up   explode
           'it exploded'

Ngalakan, a Gunwinyguan language of northern Australia (Merlan (1983)), exhibits considerable freedom of word order, but one particle *ŋara* 'perhaps' normally occurs at the end of the sentence, as in (130).

(130)   ŋiñ-ganammup ŋara
           2SG-deaf   perhaps
           'perhaps you are deaf'

And in Lezgian, although modifiers of nouns otherwise precede the noun, the word *kwaz* 'even' follows the noun, when it is modifying a noun, as in (131).

(131)   či   Qabustanba-dikaj sew-er-iz-ni   kwaz kič'e-da
           1PL.GEN Qabustanba-CASE bear-PL-DAT-also even afraid-FUT
           'even bears are afraid of our Qabustan-ba!'

13   Examples of summaries of word order properties

In this section, we will briefly summarize the word order properties of two languages, illustrating how these languages conform or do not conform to the word order tendencies discussed in this chapter.

13.1   Siyin Chin

Consider first Siyin Chin, a Tibeto-Burman language spoken in Myanmar (Burma) (Naylor (1925)). The basic order at the clause level is SOV, as in (132).

(132)

tuazawkchīangina	[Mētē	mīhing-te]	[Kawlpī	hkuā]	a	shim kik	hī
after.that	Manipuri	person-PL	Kawlpi	village	3	attack	INDIC
		S		O		V	

'after that, the Manipuri people attacked Kawlpi village'

The word *a* '3' ('third person') immediately preceding the verb in (132) is a subject clitic pronoun inflecting for the person of the subject, like an agreement affix, but a separate word. It always immediately precedes the verb and is obligatory in all clauses, except in certain well-defined cases, like imperative clauses. In fact, it is not entirely clear that these pronominal morphemes are not prefixes rather than separate words; they are written as separate words by Naylor (1925), but the date and nature of the description make this questionable. Clauses in which the subject is represented entirely by this clitic pronoun have the appearance of being OSV, as in (133).

(133)   suang atam   ke   dē   hī
         stone   many 1   want INDIC
         O   'S'   V
         'I want many stones'

However, the subject clitic *ke* '1' ('first person') does not occur in subject position, but forms a tight constituent with the verb. When independent pronouns occur, they occur in subject position, before the object, and co-occur with a subject clitic immediately preceding the verb, as in (134).

(134)   amā ching hkat a hpūk yō   hī
         3SG tree   one 3 fell PAST INDIC
         S   O   V
         'he felled a tree'

Siyin exhibits a large number of characteristics expected of it as an OV language. It employs postpositions, as in (135). This example also illustrates how adpositional phrases precede the verb in Siyin.

(135)   dimlō a   ke tām tu   hī
         Dimlo LOC 1 halt FUT INDIC
         NP   Po
         'I shall halt at Dimlo'

The genitive precedes the possessed noun, as in (136).

(136) [hīshīa ching] haw a vum hī
      this   tree   bark 3 black INDIC
      Gen   N
      'the bark of this tree is black'

The same is true for pronominal genitives (possessives), as in (137).

(137) kēma laikūng tūng   in
      1SG   pencil   pick.up IMPER
      Poss   N
      'pick up my pencil'

Relative clauses also precede the noun, as in (138).

(138) [zām ngā] pā   ke mū hī
      gong steal man 1   see INDIC
      Rel   N
      'I saw the man who stole the gong'

The order in the comparative construction is Standard-Marker-Adjective, as in (139).

(139) [hīashīa in]   sāng hīshīa a līen zaw hī
      that   house than this 3 large more INDIC
              St   M   Adj
      'this [house] is larger than that house'

Manner adverbs precede the verb, as in (140).

(140) ama amunlangina hong   pai bale
      3SG quickly   to.here go if.not
          Adv   V
      'if he does not come quickly, . . .'

The example in (140) also illustrates how adverbial subordinators come at the end of the subordinate clause, as does the example in (141).

(141) ka   anasep ke man hāngina, ke kī kom hī
      1SG   work   1 finish because 1 free   INDIC
      Clause   Subord
      'because I have finished my work, I am free'

This example also illustrates how adverbial clauses normally precede the main clause. In addition, nonverbal predicates precede the copula, as in (142), the first example with a nominal predicate, the second with a locative predicate.

(142) a. ama hkuābuīte a hī   hī
          3SG villager   3 be   INDIC
                Pred   Copula
          'he is a villager'

    b. amā in   sunga a om hī
       3SG house in   3 be INDIC
             Pred   Copula
       'he is in the house'

Siyin employs a question particle which occurs at the end of the sentence, as
in (143).

(143) sai   na kāp yō zīam?
       elephant 2 shoot PAST Q
       'did you shoot an elephant?'

Because of the general lack of morphology in Siyin, it is difficult to say
whether it is predominantly suffixing or not. There are a few suffixes, however.
There is a plural suffix, illustrated above in (132), and there is a derivational
suffix *-ina* used for forming adverbs: *damnoina* 'slowly' (cf. *damno* 'slow').
On the other hand, there is at least one element that can be analysed as a prefix:
causatives are formed by aspirating the initial consonant of the verb: *kīem* 'to
decrease, to become less' vs *hkīem* 'to cause to become less'. We can at least
say that Siyin is not inconsistent with the correlation of OV with suffixes in that
it is not predominantly prefixing.

One characteristic that is common among OV languages that Siyin lacks
is case marking distinguishing the two arguments in transitive clauses (see
example (132) above). As noted in 8.3, however, this correlation is a weak one.

Because of the lack of verbal morphology in Siyin, it is difficult to determine
whether various words are verbs. If the past tense marker *yō* in (144) is a verb,
then this conforms to the tendency for auxiliary verbs to follow the main verb
in OV languages.

(144) kōma ke vawt yō hī
       1PL   1 work PAST INDIC
                V   Aux
       'we worked'

The same is true of the negative word *ngawl* in (145):

(145) kema ke ngak ngawl tu hī
       1SG 1 wait not   FUT INDIC
                V   Neg
       'I will not wait'

A final characteristic of OV languages exhibited by Siyin is that interrogative phrases in content questions occur *in situ* rather than at the beginning of sentences, as illustrated in (146).

(146)   ama   koi lai a teang zīam
         3SG   where 3 live   Q
         'where does he live?'

In terms of word order characteristics which do not correlate with the order of verb and object, adjectives follow the noun as in (147).

(147)   mīhing hpā
         man   good
         'a good man'

Adjectival modifiers of nouns can also precede the noun, as in (148); however this structure is really a relative clause (as indicated by the pronominal subject clitic), in contrast to (147), where the adjective is directly modifying the noun.

(148)   a hpā   mīhing
         3 good man
         'a good man; a man who is good'

Numerals also follow the noun, as in (149).

(149)   mīhing htum
         man   three
         'three men'

Demonstratives can precede or follow the noun, but more often precede, as in (150).

(150)   hīshīa ching
         this   tree
         'this tree'

Finally, degree words follow adjectives, as in (151).

(151)   dū   mamā
         thirsty very
         'very thirsty'

There are various further details of word order in Siyin that can be described. For example, quantifiers meaning 'many' and 'all' are similar to numerals in following the noun, as in (149) above, and in (152).

(152)   mīhing teampō
        man   all
        N   Quant
        'all the men'

Interrogative modifiers of nouns precede the noun, however, as in (153).

(153)   bangbang   nasep na vawt zīam
        what.kind.of work 2 do   Q
        'what kind of work are you doing?'

There are also a variety of constructions involving two verbs that conform to patterns typical of OV languages, though we have not specifically discussed these above. For example, modal words for ability or obligation must follow the main verb, as in the two examples in (154).

(154)   a. ama vawt htē   hī
           3SG do   can   INDIC
                    V   Modal
           'he can do it'

        b. ama vawt tu nī hī
           3SG do   must   INDIC
                    V   Modal
           'he must do it'

The word meaning 'want' follows the verb denoting what is wanted, as in (155).

(155)   a naupā   a pai nuap hāngina, . . .
        3 younger.brother 3 go   want because . . .
                            V   Want
        'because his younger brother wanted to go, . . .'

And expressions of purpose precede the main verb, as in (156).

(156)   [ngasā shia natu] yingtung-tunga   ke pai nuam hī
        fish   fish PURP early.in.morning 1 go want INDIC
        Purp   V
        'I want to go out early in the morning to fish'

13.2   Batad Ifugao

The second language whose word order we will give an overview of is Batad Ifugao, an Austronesian language spoken in the Philippines. Batad Ifugao is a verb-initial language, but, before we can discuss the word order, it is necessary to discuss briefly some basic features of the grammar of this language. Batad

Ifugao is like other Philippine languages (e.g. Tagalog, Cebuano) in having what is traditionally called a *focus system*, where one nominal in each clause has the privileged status of being the grammatical topic and the verb inflects for the semantic or grammatical relation of this nominal to the verb. Noun phrases occur with determiners which code a variety of grammatical properties and which interact with the focus system. The focus system in Batad Ifugao is in some respects more complex than that found in other Philippine languages, and, because the primary concern here is to illustrate word order in this language, the glosses for most of the examples cited below do not include information about the focus form of the verb, and determiners are simply glossed 'DET'. The nature of the focus system in Batad Ifugao raises questions about what should be called 'subject' (cf. Schachter (1976, 1977) for discussion of the issues surrounding this question for Tagalog), and thus presents a question of how to classify the language according to the typology of SVO, VSO, etc. However, the verb normally precedes its arguments, regardless of the clause type, as in (157), so that even in the absence of identifying an element as subject, the language is clearly verb-initial.

(157)   a. ginumhob hi   Manābung hi   āyiw
            PAST.burn DET Manabung DET wood
            'Manabung burned some wood'

        b. in-dat   Aligūyun nan dotag ay   agī-na
            PAST-give Aligūyun DET meat   DET brother-3SG.POSS
            'Aligūyun gave the meat to his brother'

In fact, the actor nominal immediately follows the verb, regardless of whether it is topic, so that if we ignore the issue of topic, and assume that the actor is the subject, the language might be characterized as VSO.

    As in most verb-initial languages, nonverbal predicates also occur at the beginning of the clause. This is illustrated for a nominal predicate in (158a) and for an adjectival predicate in (158b).

(158)   a. binabāi   nin   di   denngol-mu
            woman.PL perhaps DET hear-2SG
            'perhaps those whom you heard were women'

        b. adangyan hi   Habbēleng
            rich   DET Habbēleng
            'Habbeleng is rich'

    Batad Ifugao conforms fairly closely to word order characteristics associated with verb-initial languages in particular and with VO languages in general. The genitive follows the noun, as in (159).

(159)   di   payaw   Wīgan
        DET   pond.field Wigan
              N   Gen
        'Wigan's pond field'

When the word preceding the genitive noun phrase ends in a consonant (as
payaw 'pond field' does in (159)), there is no marker of the genitive relation.
However, when the word preceding the genitive ends in a vowel, a linking
morpheme -*n* is added to this word, as in (160).

(160)   ulu-n   nan ūlog
        head-LINK DET snake
        'the head of the snake'

Pronominal genitives (possessives) also follow the possessed noun, and are
attached as enclitics, as in (161).

(161)   han imbaluy='u
        DET   child=1SG.POSS
        N   Poss
        'my child'

Relative clauses also follow the noun, as in (162).

(162)   nan baluy [an   iny-ammā-na]
        DET house LINK PAST-make-3SG
              N   Rel
        'the house that he made'

Articles, in contrast, precede the noun, as illustrated by various instances of
words glossed 'DET' in the examples, including the words *hi*, *nan*, *han*, *di*, and
ay. As discussed above, we use the term *article* in this chapter to denote words
that commonly occur in noun phrases in a language and which code various
grammatical or semantic features of the noun phrase, such as definiteness, but
do not necessarily code a definite–indefinite distinction. The system of articles
in Batad Ifugao is rather complex, and a number of the articles are used for
a range of different functions, varying with the semantic case of the nominal,
with whether it is grammatical topic or not, with whether the noun it goes with
is a proper noun or a common noun, and in some cases with pragmatic features
similar to definiteness. For example, the article *hi* can be used with a topic or
with a nontopic object, but not with a nontopic actor. As a marker of topics,
however, it only occurs with proper nouns, not with common nouns, while with
nontopic objects, it only occurs with common nouns and in that use expresses
something like indefiniteness. Example (157a) above illustrates both of these
uses of *hi.*

The order in the comparative construction in Batad Ifugao is Adjective-Marker-Standard, as in (163).

(163)

adukkoy	han	āyiw	an	lengngoh	Gumman	ya	un	nan	lengngoh	Lahhin
long	DET	tree	LINK	cut.down	Gumman	than	DET		cut.down	Lahhin
Adj						M	St			

'the tree that Gunman cut down is longer than the one that Lahhin cut down'

Adverbial subordinators occur at the beginning of their clause, as in (164).

(164)

ababāin	ahan	din	in-at	Immamata''on	ti
shameful	very	DET	PAST-do	Immamata''on	because
					Subord

in-ihdā-na-y	imbaluy-na
PAST-eat-3SG-DET	child-3SG.POSS
Clause	

'what Immamata'' on did was very shameful because he ate his son'

The example in (164) also illustrates an adverbial subordinate clause following the main clause. Both orders of main and subordinate clause are in fact common in Batad Ifugao, depending in part on the type of clause; (165) illustrates an adverbial clause preceding the main clause.

(165)

wa	an	tuma'dog	hi	Pīlay	ya	in-lumdit	Pū'it	di	hu'ī-na
when		stand.up	DET	Pīlay	and.then	PAST-tramp	Pū'it	DET	foot-3SG.POSS

'when Pilay stood up, Pu'it tramped on his foot'

Batad Ifugao employs a particle for marking polar questions. While the short form of the particle can occur in various positions in a sentence, the long form *undan* occurs at the beginning of the sentence, as in (166), conforming to the typical pattern for verb-initial languages.

(166)

undan	lumagalāga	hi	Bukkāhan	hi	ulbung
Q	weave.HABIT	DET	Bukkahan	DET	rice.basket

'does Bukkahan always weave rice baskets?'

Interrogative expressions in content questions in Batad Ifugao occur at the beginning of the sentence, as in (167).

(167)

angnganggoh	di	pohdo-m	hi	ono-m
what	DET	want-2SG	DET	eat-2SG

'what do you want to eat?'

(literally 'what is that which you want to eat?')

Batad Ifugao does not have many adpositions, since the role played by adpositions in other languages is carried partly by the verb morphology that is involved in the focus system and partly by the form of determiners, and the adpositions that do exist do not appear to play a major role in the language, but what adpositions exist are prepositions rather than postpositions, as in (168) – the one in (168a) meaning 'before', the one in (168b) meaning 'like'.

(168)   a. mahhūna' an   um-uy ya un he''a
           be.first   LINK go   before 2SG
                                           Pr   NP
           'I will go before you'

b. ay   ugāli-n   di   i-Batad   di   aton nan iy-Umbūlu
   like custom-LINK DET from-Batad DET do   DET from-Cambulo
   Pr   NP
   'what the people of Cambulo do is like the custom of the people of Batad'

Prepositional phrases normally follow the verb in Batad Ifugao, as in (168a). The prepositional phrase in (168b) occurs at the beginning of the clause because it is itself the predicate and is not modifying a verb. Its initial position reflects the position of predicates in general in Batad Ifugao, whether they are verbal or nonverbal.

In the preceding section on Siyin Chin, we noted that modal words for ability follow the main verb. In Batad Ifugao, we find the opposite order: the verb meaning 'be able' precedes the verb it goes with, as in (169).

(169)   mabalin-a' an   dalānon nan adagwi-n kulha
        able-1SG   LINK walk   DET far-LINK road
        $Verb_1$-S   Link V_2
        'I am able to walk far on a road'

The construction in (169), what we can call a 'verb chain construction', is one that is used for a wide range of meanings in Batad Ifugao: it consists of a pair of verbs, connected by the linking word *an*. Typically the subject will immediately follow the first verb. It is not clear that there is any sense in which one of the two verbs is grammatically the 'main verb', though semantically, the first verb is often one that one might loosely call an 'auxiliary verb', while the second verb is more like a main verb. Only a minority of verbs in the language can occur as the first verb in this construction, while apparently any verb can occur as the second verb. A few other examples are given in (170), where the first verbs mean 'want', 'begin', and 'delay until night', respectively. (The example in (170a) is embedded within a nominal expression, and there is no overt subject.)

(170) a. di   [mamhod   an   manāyaw]
          DET   want   LINK   dance
          V₁   Link   V₂
          'the ones who want to dance'

       b. ente''an da Umāngob an   mumbā'i
          begin   PL Umangob LINK   recite.ritual.prayer
          V₁   S   Link   V₂
          'Umāngob and others began reciting ritual prayers'

       c. iny-ahdom   Tumāpang   an   iyanāmut   din   lāman . . .
          delay.until.night Tumapang   LINK   bring.home that wild.pig
          V₁   S   Link   V₂
          'Tumapang delayed until night bringing home that wild pig . . .'

Batad Ifugao is typical among verb-initial languages in placing the expression
of words meaning 'want' and 'begin' before the verb expressing what is wanted
or what has begun.

   This verb chain construction is also used to express a number of meanings
that are often expressed in other languages by adverbs. For example, in (171),
the first verb in the verb chain means 'to be first' and when combined with a
verb, the resultant verb chain has the meaning 'to be the first to "verb"'.

(171) mahhūn-a'   an   umuy ya un he''a
       be.first-1SG LINK   go   before 2SG
       V₁-S   Link   V₂
       'I will go first before you'

Batad Ifugao uses this construction to express meanings that other languages
express with manner adverbs, where the first verb is a verb expressing manner
and the second verb is a verb expressing the action that is done in the manner
expressed by the first verb, as in (172).

(172) a. umulla'ullay   han nundogoh an   dumālan
          do.slow:HABIT   DET be.in.pain LINK walk
          V₁   S   Link   V₂
          'the one who is in pain always walks slowly'

       b. imay'an-yu-n   uminum hinan bayah
          do.moderately-2PL-LINK   drink DET   homemade.beer
          V₁   -S-Link   V₂
          'drink rice beer moderately'

From a purely semantic point of view, we might say that Batad Ifugao places the
manner adverb before the verb, contrary to the normal pattern for verb-initial

languages. But this characterization would be misleading, since it is not at all clear that the verb expressing manner is *modifying* the other verb, rather than the two verbs simply forming a chain in which neither verb is modifying the other. The situation here is similar to the one discussed above for Jakaltek and illustrated in (47), where the manner expression is the main verb and the other verb is subordinate to it. While there is no evidence in Batad Ifugao that the other verb is subordinate to the verb expressing manner, there is no evidence that the verb expressing manner is in any sense modifying the other verb, and hence we can say that Batad Ifugao simply does not have manner adverbs that modify verbs, and hence the construction in (172) is not relevant to universals regarding the order of manner adverb and verb.

A version of the verb chain construction is used for expressions of purpose, as in (173).

(173)
immuy hi   Bumallātung hi ad Bannāwol an   mungngīna-h bābuy
go   DET   Bumallātung DET Bannāwol LINK buy-DET   pig
V   Purp
'Bumallātung went to Bannāwol to buy a pig'

This construction differs from the verb chain construction in that the second verb follows a locative expression; in the verb chain construction, only the subject can occur between the two verbs. This example also illustrates the fact that purpose expressions normally follow the main verb in Batad Ifugao.

Let us turn now to the order of various elements whose order does not correlate with the order of verb and object and for which we therefore have no expectations of what to find in Batad Ifugao on the basis of its being verb-initial. We saw earlier that genitives and relative clauses follow the noun in Batad Ifugao while articles precede. It turns out that various other modifiers of nouns, including demonstratives, numerals, and adjectives, precede the noun, as illustrated in (174).

(174)   a. din   apuy
            that   fire
            Dem   N
            'that fire'

         b. nan tulu-n   balāhang
            DET three-LINK young.girl
                 Num   N
            'the three young ladies'

c. hinan ongol an   batu
   DET   large LINK stone
           Adj   N
   'a large stone'

Note that in both (174b) and (174c), with a numeral and with an adjective modifying a noun, there is a linking morpheme, either the enclitic *-n* attached to the numeral in (174b) or the separate linking word *an* between the adjective and the noun in (174c). We have seen this same linking morpheme in various other constructions in this language, in the genitive construction in (160), introducing the relative clause construction in (162), preceding the second verb in the verb chain construction in (169) to (172), and preceding a purpose expression in (173).

While single adjectives modifying a noun precede the noun, as in (174c), adjective phrases containing an additional word modifying the noun normally follow the noun. Contrast (175a), with a single adjective modifying the noun and preceding the noun, with (175b), where the adjective is modified by a verbal expression *panniga'* 'with reference to the way I see it' and the adjective plus modifier now follow the noun instead of preceding it.

(175)   a. nan nappuhi-n tibung
          DET bad-LINK   wine.jar
              Adj   N
          'the bad wine jar'

    b. nan tibung   an   [nappūhih   pan-nig-a']
          DET wine.jar LINK bad   MANNER-see-1SG
              N   Adj+Modifier
          'the wine jar which is bad with reference to the way I see it'

The normal means of expressing degree with adjectives is morphological, but there is a degree word *ahan* 'very', which follows the adjective, as in (176).

(176)   ababāin   ahan
        shameful very
        'very shameful'

Negation is expressed in Batad Ifugao by a word that occurs at the beginning of the sentence, preceding the verb. It is not clear whether the negative word is itself a verb. It does behave verbally to the extent that there are two negative words used in verbal clauses, with a difference in tense, as illustrated in (177), with the past tense negative in (177a) and the nonpast negative in (177b).

(177) a. agguy   nolo' han imbaluy-'u
          NEG.PAST sleep DET child-1SG.POSS
          'my child didn't sleep'

       b. adī-da   um-uy
          NEG.NONPAST-3PL go
          'they will not go'

While verbs also distinguish past and nonpast forms (not indicated in the glosses in this chapter), the relationship between the two negative words is suppletive and these words exhibit no other morphological properties of verbs, so it is not clear whether they ought to be considered verbs. They do serve as host to subject enclitic pronouns, like *-da* 'third plural' in (177b), a property that is also one shared by verbs, although this property is also shared by a repetitive aspect particle *gun*, as in (178).

(178)   gun-na   inhanglag din gadiw
        repeatedly-3SG roast   that fish(sp.)
        T/A   V
        'he repeatedly roasted those small river fish'

Since this aspect particle does not otherwise exhibit any verbal properties, it is likely that the ability to host subject enclitics is a function of the fact that the negative words and this aspect particle can occur in initial position in the clause and that the ability to host these clitics reflects nothing more than this. Note that if we consider the negative words to be verbal, i.e. negative auxiliaries, then their position before the main verb is what we would expect of a negative auxiliary in a verb-initial language. However, negative particles also tend to precede verbs, as described above in section 7.4, regardless of the order of verb and object, and this tendency seems to be particularly strong in verb-initial languages, so the position of negative words in Batad Ifugao is unsurprising, regardless of whether the negative is verbal or not.

   The word *gun* 'repeatedly', illustrated above in (178), is a candidate for status as a nonverbal tense–aspect particle, and it precedes the verb. Most indications of tense and aspect in Batad Ifugao are expressed in the verb morphology. Another apparent nonverbal tense–aspect particle that precedes the verb is one that is used, along with reduplication of the first syllable of the verb, to indicate that several events are happening concurrently, as in (179).

(179)   'ahi   ga-gallaw   nan linalā'i-n   mangngal hi   dotag
        CONCUR CONCUR-crowd DET man.PL-LINK get   DET meat
        'the men all crowd together in getting meat'

   It is noted above in 8.2 that prefixing is more common in VO languages than it is in OV languages. Batad Ifugao conforms to this in that it has a moderate

amount of prefixing, and it is VO. It also employs a fair amount of suffixing and infixing. The use of infixing is more common in VO languages than it is in OV languages, so again this fits the fact that Batad Ifugao is verb-initial. Because the glossing in the examples above generally did not give the details of verb morphology, we need to examine a few examples illustrating the different types of affixes. The focus affixes indicating the semantic category of the grammatical topic is often a prefix, as illustrated by the actor focus prefix in (180a); it is sometimes a suffix, as illustrated by the locative focus suffix in (180b); and it is sometimes an infix, as illustrated by the actor infix *-um-* in (180c) (the verb stem is *hagīlip* 'spin horizontally'; the initial *h-* is part of the stem).

(180)   a. umuy mang-anup nan linalā'i
             go   ACT-hunt   DET man.PL
             'the men will go to hunt'

        b. pagūy-an nan tatagud   Nāyun nan ūma-da
            plant-LOC DET person.PL Nayon DET upland.field-3PL.POSS
            'the people of Nayon plant their upland fields with rice'

        c. h<um>agīlip   nan batu
            <ACT> spin.horizontally DET stone
            'the stone will spin horizontally through the air'

The example in (181) illustrates two infixes, a past tense infix *-in-* and an actor focus infix *-um-*. (The verb stem for 'rain heavily' without infixes is *doloh*.)

(181)   d<in><um>loh   ad   a''u
        <PAST><ACT>rain.heavily DET.LOC last.night
        'it rained heavily last night'

All three types of affixing are common in the language. This is a normal pattern for a verb-initial language.

14   Summary

We have discussed a variety of different pairs of elements in this chapter, some of which exhibit correlations with the order of object and verb and some of which do not. The following chart summarizes these generalizations, ignoring the fact that some are unidirectional and ignoring the fact that, for some characteristics, SVO languages exhibit a pattern intermediate between OV and verb-initial:

OV	VO
postpositions	prepositions
genitive–noun	noun–genitive
manner adverb – verb	verb – manner adverb
standard–marker	marker – standard
standard–adjective	adjective – standard
final adverbial subordinator	initial adverbial subordinator
adpositional phrase – verb	verb – adpositional phrase
main verb – auxiliary verb	auxiliary verb – main verb
predicate – copula	copula–predicate
final question particle	initial question particle
final complementizer	initial complementizer
noun–article	article–noun
subordinate clause – main clause	main clause – subordinate clause
relative clause – noun	noun – relative clause
noun – plural word	plural word – noun

The following list of pairs of elements are those whose order does not correlate with that of object and verb:

adjective, noun
demonstrative, noun
numeral, noun
negative particle, verb
tense–aspect particle, verb
degree word, adjective

15   Suggestions for further reading

The classic work in word order typology is Greenberg (1963) (sometimes cited as Greenberg (1966), its apparently unrevised second edition). This work not only documents many of the patterns that correlate with the order of object and verb, but is often viewed as defining the beginning of the modern study of linguistic typology in general. Hawkins (1983) provides a detailed discussion of various aspects of word order typology. Evidence supporting many of the claims made in this chapter is given in Dryer (1992). In this chapter, we have avoided discussion of the question *why* the order of various pairs of elements correlates with that of object and verb, but there is an extensive literature on this topic, including Aristar (1991), Dryer (1992), Frazier (1979), Givón (1975, 1984a), Hawkins (1983, 1984, 1990, 1994), Keenan (1979), Kuno (1974), Lehmann (1973, 1978a), Vennemann (1973, 1974a, 1974b, 1976), and Vennemann and Harlow (1977). A wide variety of different explanations have been proposed, some in terms of syntax, some in terms of semantics, some in terms of sentence

processing, some in terms of grammaticization, and even some in terms of phonology. In addition to the sources mentioned above, there are many references in the generative literature to a distinction between head-initial and head-final languages, which assumes an explanation for all or part of the correlations. See Dryer (1992) for problems with this view. It should also be noted that some of the literature discussing correlations or proposing explanations for the correlations assumes some correlations that can be shown not to be correct (see Dryer (1992)), often assuming incorrectly, for example, that the order of adjective and noun correlates with the order of object and verb.

Further information on word order, with accompanying maps that show the geographical distribution of different types, is given in Dryer (2005) and sixteen other chapters on word order by Dryer in Haspelmath *et al*. (2005).

3 The major functions of the noun phrase

Avery D. Andrews

0 Introduction

In this chapter we will discuss the major functions of noun phrases (NPs) in the languages of the world. We can think of NPs as having three different kinds of functions: semantic, pragmatic and grammatical. Semantic and pragmatic functions are aspects of the meanings of sentences, grammatical functions aspects of their structure.

Semantic functions, often called semantic roles, are the different ways in which a sentence can describe an entity as participating in a situation. Consider (1):

(1)   The farmer kills the duckling

Here the verb *kill* indicates that we have a situation in which one entity kills another. It provides two semantic roles, 'killer' and 'killed', taken by the referents of the preverbal NP *the farmer* and the postverbal NP *the duckling*, respectively. In order for the sentence to be true, the entities referred to by these NPs must act or be acted upon in accord with these roles. Semantic roles are thus an aspect of the relation between sentences and the situations they refer to.

But language is used not merely to depict the world, but to communicate in it: its users are part of the world they talk about. There is therefore a further aspect of meaning, concerning more than than just what a sentence is about, which contributes to determining when it may be used. This aspect of meaning, called pragmatics, involves such things as the hearer's presumed ignorance or knowledge of various features of the situation being talked about, the presumed spatial and social relationships between the speaker and the hearer, what the speaker thinks the hearer might be attending to, what the speaker wants the hearer to take special notice of, and so forth. These constitute ways in which utterances with the same objective content can fulfil different communicative purposes. Properties of NP that relate the sentence to its context of use without affecting objective content are called pragmatic functions.

In English, for example, (1) has the variants shown in (2):

(2)   a. It is the farmer that kills the duckling
          b. It is the duckling that the farmer kills

The sentences of (2) designate precisely the same kind of situation as (1). But (2a) presumes that the hearer knows that somebody or something kills the duckling, but not who or what; and (2b) presumes that the hearer knows that the farmer killed somebody or something, but not who or what. Example (1), on the other hand, in its most straightforward articulation, with neutral intonation, does not presume that the hearer knows anything about the event of killing. These sentences therefore give their NPs the same semantic roles, but different pragmatic functions. We will say that (2a) 'focuses' on the killer of *kill* (treating it as new information and as the unique entity filling the role of killer), and that (2b) does the same thing for the role of entity killed.

The semantic roles and pragmatic functions of the NPs in a sentence may be called their 'semiotic functions', since they have to do with the meaning of the sentence. Semiotic functions are ultimately signalled by 'overt coding features' such as word order, case marking and cross-referencing (agreement). But it is usually quite difficult to provide a coherent account of how this occurs in terms of a direct connection between the coding features and the semiotic functions they express. Rather it normally seems better to posit an intervening level of 'grammatical structure': the coding features indicate the grammatical structure of the sentence, and the grammatical structure determines the semiotic functions.

The grammatical functions of NPs are the relationships in this grammatical structure which matter for determining the semantic roles and grammatical behaviour of NPs. For example, in (1) we recognize the grammatical functions of 'subject' (preverbal NP) and 'object' (postverbal NP). There is a rule for using the verb *kill* which says that the subject should express the 'killer' role and the object the 'killed' role. The semantic role of an NP is thus determined jointly by the verb and the grammatical function of the NP. The structural positions of *the farmer* and *the duckling*, of (2a) and (2b) respectively, likewise cause them to have the pragmatic function of focus.

Grammatical functions are also important for principles governing the form of sentence structure. A familiar example is the principle of subject–verb agreement in English, whereby a present-tense verb with a third person singular subject takes a special form ending in /-z/. Thus, if the subject of (1) is pluralized, the form of the verb must change, but pluralizing the object does not have this effect:[1]

---

[1] In these and subsequent examples, '*' within parentheses indicates that the example is bad if the material within the parentheses is included, while '*' immediately in front of them indicates the example is bad if the material within the parentheses is omitted. Hence the verbal ending is impossible in (3a), but obligatory in (3b).
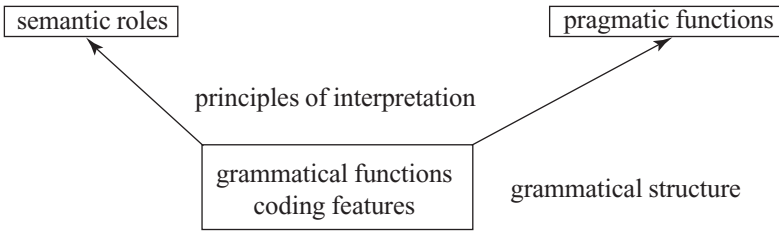
Figure 3.1 Organization of grammatical structure

(3)   a. The farmers kill(*s) the duckling
        b. The farmer kill*(s) the ducklings

The grammatical function of subject is thus involved in this constraint on the form of English sentences.

The relationships between semiotic functions, grammatical functions and coding features may be illustrated as in figure 3.1: principles of grammatical structure determine the distribution of grammatical functions and how they are expressed by coding features; on the basis of grammatical structure, principles of semantic interpretation determine the assignment of semantic roles, pragmatic functions and other aspects of meaning not considered in this chapter, such as logical scope of quantifiers.

Here we will be primarily concerned with the grammatical functions of NPs in clause-structure. But since the task of the grammatical functions is to express the semantic and pragmatic ones, we first need to survey these briefly. This will be done in section 1 of the chapter, where the coding features will also be discussed. Then in section 2 we will present a basic classification of grammatical functions into three types, 'core', 'oblique' and 'external', and discuss the latter two. In section 3, we will discuss the core grammatical functions in detail. Finally, in section 4, we will discuss phenomena that suggest a re-evaluation of the standard view of grammatical relations, at least for some languages.

In the literature, the term 'grammatical relation' is used as a virtual synonym of 'grammatical function'. However, we will find it useful to differentiate between these terms here. A 'grammatical function' will be any definable relationship which it might be useful to recognize in the sentence structures of a language, regardless of how important it seems to be, or how sensible it might be to see it as a *primitive* ingredient of sentence-structure. A 'grammatical relation' on the other hand will be a grammatical function that is of particular importance for the workings of the language, so that it would be reasonable, although not necessarily correct, to regard it as a primitive ingredient of sentence-structure. This terminological distinction, although novel, is useful for discussing sentence-structures in a language without making controversial claims about what their ultimate analysis ought to be, and what kind of linguistic theory they ought to be framed in.

For example, in English, subject and object are grammatical relations, since they are relevant for the operation of many grammatical rules, so that one could plausibly view them as primitive ingredients of English sentence-structure. But 'subject of a transitive clause' ('transitive subject') and 'subject of an intransitive clause' ('instransitive subject'), although they are grammatical functions (since they are definable within any reasonable theory of English sentence structure), do not qualify as grammatical relations in English, since they are not relationships that are relevant for the operation of a significant number of grammatical rules, and treating them as primitives of sentence-structures will obscure the statement of grammatical rules (one would have to say 'verbs agree with their transitive subjects or their intransitive subjects, whichever is present', rather than just 'verbs agree with their subjects').

1   Preliminaries

1.1   Semantic roles

In the most usual type of sentence structure, there is a verbal element that designates a type of situation, which usually implies various roles, that is, ways of participating in that situation. Thus we have seen that *kill* designates a type of situation with 'killed' and 'killer' roles, among others. The element that defines the type of situation and the roles we call a 'predicate',[2] the NPs filling the roles we call 'arguments'.

The predicate needn't be a single verb. Sometimes it is a complex consisting of several verbs, or a verb plus a nominal or adverbial element. Example (4a) illustrates a two-verb predicate from the Papuan language Barai (Foley and Olson (1985)), (4b) a verb + noun predicate from the Dravidian language Malayalam (Mohanan (1982)), and (4c) a verb + (adverbial) particle predicate from English. The complex predicate in each example is italicized: complex predicates have recently been the subject of a great deal of research; see Alsina, Bresnan, and Sells (1997) for a recent collection of studies.

(4)   a. Fu fase isema   *fi   isoe*
          he letter wrongly sat write
          'He wrongly sat writing a letter'

       b. Kuṭṭi ammaye *salyam   ceytu*
          child mother annoyance did
          'The child annoyed the mother'

       c. The guards *beat* the prisoners *up*

---

[2] Note that this is different from the use of the term 'predicate' in traditional grammar to refer to the verb and its objects and complement.

Languages also have sentence types in which a nonverbal element is the predicate, or where there is no overt predicate word, the predicate being understood from the syntactic structure of the sentence as a whole. We illustrate this possibility with some examples from Russian (see section 1 of chapter 4 by Dryer, for more discussion):

(5)   a. Kniga na stole
            book   on table
            'The book is on the table'

       b. U menja kniga
            of me   book
            'I have a book'

In addition to a main predicate, a sentence may have additional, subsidiary predicates. In the sentence *John made Mary happy*, for example, the principal predicate is the verb *made*, and the adjective *happy* is a subsidiary predicate applying to *Mary*. In spite of these possibilities, we will generally refer to the main predicate simply as 'the verb'.

A predicate defines a set of highly specific roles, such as 'killer' and 'killed', which can in fact be thought of as being rather like roles for the actors in a drama: the role determines what happens to its filler. Examining the nature of the relation between these roles and grammatical relations, we find that it is far from arbitrary: there are always far-reaching regularities and generalizations, statable in terms of semantically definable classes of roles. Thus it is no accident that *kill* expresses the killer as subject and the killed as object; *kill* is one of a large class of verbs in which one participant, possibly exercising his or her will, does something to another which significantly affects the other. When two-participant verbs in English meeting this description are in their active form (we will discuss passives later), they always have the acting, 'agent' argument as subject, and the acted-upon or 'patient' argument as object.

I will use the term 'semantic role' to refer to both the specific roles imposed on NPs by a given predicate, such as 'killer' and 'killed', and to the more general classes of roles, such as 'agent' and 'patient'. Semantic roles are important in the study of grammatical functions since grammatical functions usually express semantic roles in a highly systematic way.[3] In our subsequent discussion we will first examine the agent and patient roles, and the intimate connection they have with the basic grammatical forms of all languages. Then we will survey a variety of further semantic roles which it is useful to recognize.

---

[3] Semantic roles first began to be discussed extensively in recent American linguistics in the work of Gruber (1965, 1976) and Fillmore (1968). For more recent discussion, see, for example, Jackendoff (1990), Dowty (1991) and Wechsler (1995).

1.1.1 Agent and patient To begin with, it is essential to understand that there is an element of arbitrariness in the definitions of agent, patient or any other semantic roles. We try to define them in such a way that they will be most useful for helping us to identify and understand phenomena, but there will always be issues that people can disagree about. For example some people might think that agents should be conscious and volitional performers of their actions; others might be happy with unconscious and accidental agency. Paradoxically, it's the very importance of these concepts that makes it difficult to be sure about the best way of defining them: the fully volitional performer of an action, and the substantially affected undergoer of one, seem to be 'grammatical poles' in the sense that other semantic roles that don't quite meet these criteria, such as the Seer and Seen of the verb *see*, tend to be expressed in the same way. The assimilation in mode of expression of many different semantic roles to agent- and patient-like concepts makes it hard to work out how best to define these concepts.[4]

Another point is that we need to distinguish between what a verb itself actually implies, and what might be true in a situation described by the verbs. In a situation described by the sentence 'Mary hit John', for example, Mary might intend to hit John, or hit him by accident. The sentence itself is neutral on this issue. In our accounts of semantic roles, what we will be interested in is what the verbs and sentences themselves imply, not what is actually the case in the situation described.

With these cautions in mind, I will define an agent as a participant which the meaning of the verb describes as doing something, or causing something to happen, possibly intentionally (that is, because (s)he wants it to). We take intentionality as a possible but not required property of the role because in many languages, such as English, many verbs, such as *hit* as discussed above, are neutral about intentionality. On the other hand, if a language has constructions in which the causer of an action is explicitly characterized as not intending it, such as the 'Involitive' forms of Singhala (Inman (1993)) and many other South Asian languages (Klaiman (1986)), these causers will not be classified as agents.

A patient will be defined as a participant which the verb describes as having something happen to it, and as being affected by what happens to it. By this definition, the objects of *kill*, *eat* and *smash* are clearly patients, while those of *watch*, *hear*, and *love* are clearly not. The objects of *hit* and *kick* are intermediate in status, because although something obviously happens to them, they are less clearly affected by it. In most languages, NPs with these roles behave like patients, and can be considered as marginal instances of this role.

---

[4] See Dowty (1991) for a very useful discussion of this problem.

But sometimes their grammar is significantly different. For example in Northwest Caucasian languages such as Abkhaz and Adyghe, verbs with meanings such as *beat*, *stab*, and *push*, which we would tend to think of as taking patients, take a different case-marking pattern from verbs with meanings such as *kill*, *write*, or *see*, illustrated here with examples from Adyghe (Catford (1976:44), see the beginning of the volume for an explanatory list of abbreviations used in the glosses):

(6)   a. bojetsɪ-m   pɪjɪ-r   ɪwɪk'ʀ
           warrior-ERG enemy-NOM killed
           'The warrior killed the enemy'

       b. bojetsɪ-r   pɪjɪ-m   jepɪdʒʀ
           warrior-NOM enemy-ERG stabbed
           'The warrior stabbed the enemy'

The *stab*-type verbs are taking the same case-marking pattern as verbs taking non-patient arguments, with meanings such as 'help' and 'wait-for', which frequently diverge from the standard treatment of full patients. The examples indicate that the ERG-NOM pattern is used when the patient changes its state, the NOM-ERG pattern when it doesn't.

Agent and patient play a fundamental role in all languages. The class of two-argument verbs taking an agent and a patient is important enough to give it a name: we shall call these verbs 'primary transitive verbs' (PTVs). Languages always seem to have a standard way or small set of ways in which they normally express the agent and patient of a PTV. If an NP is serving as an argument of a two-argument verb, and receiving a morphological and syntactic treatment normally accorded to an agent of a PTV, we shall say that it has the grammatical function A; if it is an argument of a verb with two or more arguments receiving a treatment normally accorded to the patient of a PTV, we shall say that it has the grammatical function P.[5] Abkhaz and Adyghe, as illustrated above, are unusually limited in the extent to which they extend the grammatical treatment of PTVs to verbs that don't have the core semantics of PTVs. It is a further unusual feature of these languages that the same case form is used for the agent of PTVs as for the more patient-like argument of two-argument non-PTVs. Two-argument non-PTVs with significant difference in appearance from PTVs are frequently called 'semi-transitive'; for further discussion of semi-transitives, see Dryer, chapter 4, section 2.5.

It is especially important to emphasize that we are speaking of the grammatical treatment associated with the semantic roles, not the semantic roles

---

[5] A widely used alternative to P is the label O, which is in fact the original notation for the concept, introduced in Dixon (1972:xxii). In conformity with the other chapters in this volume, we here use P to indicate the affiliation of the syntactic concept with the semantic role of patient, in the same way that A reflects the affiliation with agent.

themselves. In an English sentence such as *John likes Mary*, John is not an agent, and Mary is not a patient, but John is an A and Mary is a P, because these NPs are getting the same grammatical treatment as an agent and a patient of a PTV.

A sentence is called 'transitive' if it has A and P functions in its syntactic structure, 'intransitive' if one or both of these is missing. These definitions apply to the possibly abstract syntactic structure of the sentence: the NPs needn't appear in the overt, visible form. An NP in an intransitive sentence that is receiving the treatment normally accorded to the single argument of a one-argument predicate will be said to have S function. Languages always seem to have A and P functions, in the sense of having a uniform treatment of agent and patient of a PTV. On the other hand we will see in section (4.3.2) that it may be the case S is sometimes absent.

A, S and P are important because languages always seem to use PTVs as a grammatical model for a great many other types of verbs. We have already mentioned *like* as a verb that takes non-agent A and non-patient P, and there are many more. *See*, for example, is like this in most languages, while the Liker and Liked of *like* are often expressed differently from agent and patient. The widespread use of PTVs as a syntactic model makes it difficult to be absolutely precise about drawing the boundaries of the class, but, fortunately, a high degree of precision is not required.

A, S and P are grammatical functions, not grammatical relations, though often one of them coincides with a grammatical relation in a language. In English, for example, P can be identified with the grammatical relation 'object', but neither A nor S by themselves can be identified with 'subject', since A comprises transitive subjects and S intransitive ones, neither of which are plausible grammatical primitives of English sentence-structure, because too many principles of English grammar would have to be formulated in terms of A or S individually. But they are grammatical functions, because they are easily definable in terms of any set of plausible primitives for English sentence structure, for example A as 'subject of a sentence that has an object', and S as 'subject of a sentence that does not have an object'.

Although A, S and P cannot in general be regarded as grammatical relations, they are closely related to them, and they are furthermore associated with the syntactically most active ones, those most important in the grammatical system of a language. Hence identifying them is the first step in working out the system of grammatical relations in a language.

Most often, one finds one grammatical relation associated with A and S, and another with P. The former can be called a 'canonical subject', the latter a 'canonical object'. But as we shall discuss below, there are a number of languages in which canonical subjects and objects don't exist. For such languages, there is usually a debate about whether the terms 'subject' and 'object' should

be used at all, and, if so, what they should be applied to. In this chapter, 'subject' and 'object' will therefore be taken as recurrently convenient terms, rather than presumedly universal grammatical primitives.

1.1.2   Other semantic roles   Besides agent and patient, a number of other semantic roles are also important for grammar. Semantic roles in general may be divided into two rough classes: participatory and circumstantial. Participatory roles are borne by what one would think of as actual participants in the situation implied by the verb. Agent and patient are the most essential and typical participatory roles. Circumstantial roles are borne by entities that do not really participate, but instead form part of the setting of the event. Benefactive, the person for whom something is done, is a typical circumstantial role.

   Aside from agent and patient, some of the other more important participatory roles are: directional, with source and goal subtypes; 'inner' locative (giving the location of a participant, rather than of the event or state as a whole); experiencer (a participant who is characterized as aware of something); recipient (a participant who 'gets' something); theme (a participant which is characterized as being in a state or position, or changing its state or position, sometimes treated as a kind of patient); causer (a participant who causes something to happen, but does not act intentionally); and instrumental (a participant that the agent uses to act on the patient). Note that the theme and patient roles are closely related, though not identical: unlike patients, themes needn't be acted upon by anything, and it is sometimes appropriate to regard as patients certain arguments, such as things that are hit or kicked, which may be regarded as affected by what is done to them, but do not necessarily undergo a clearcut change of state.

   Our list of roles is furthermore not supposed to be a valid and thorough classification of all forms of participation, but simply an assortment of ones which get distinctive treatment by grammars often enough to be worth setting up names for. Here are some examples of these roles:

(7)   a. Tiger snakes$_{theme}$ inhabit Australia$_{innerlocative}$
       b. George$_{agent\&theme}$ walked from/to the store$_{source/goal}$
       c. I$_{experiencer}$ love Lucy
       d. Frederika$_{causer}$ annoys me$_{experiencer}$
       e. Darlene$_{agent}$ handed Bruce$_{recipient}$ a sausage$_{theme}$
       f. Bill$_{agent}$ prodded the snake$_{patient}$ with a stick$_{instrumental}$
       g. The Earth$_{causer}$ attracts the moon$_{theme}$
       h. The car$_{theme}$ is expensive

Note that not every NP in these examples is labelled with one of our semantic roles. This is because no presently known system of semantic roles can be applied in a comprehensive and convincing manner. For example *Lucy* in (7c)

isn't subscripted for a role; some possibilities might be 'goal' or 'object of emotion', but no specific proposal has received widespread acceptance.

Aside from benefactive, some other important circumstantial roles are 'outer' locative, (the place where something is done), reason (why something is done), circumstantial comitative (something that accompanies a participant, but does not itself participate), and temporal. These are illustrated below:

(8)   a. Susan caught a lizard in the garden$_{outerlocative}$
         b. Bruce barbecued a sausage for Darlene$_{benefactive}$
         c. Alvin shot up a sign for fun$_{reason}$
         d. Shirley went diving with a speargun$_{circumstantialcomitative}$
         e. Jack ate a sausage during the race$_{temporal}$

The distinction between participatory and circumstantial roles is closely related to a distinction between 'arguments' and 'adjuncts' that will be introduced in section 2.3.

There are of course many (perhaps infinitely many) more semantic roles that might be significant for the grammar of a language. The ones discussed here are merely some of the more recurrent ones. It should also be pointed out that, in accord with most of the literature, we have paid no serious attention to the problem of *defining* the semantic roles, but just contented ourselves with rather vague characterizations.

1.2   Coding strategies

There are three basic techniques which languages use to code syntactic functions: order and arrangement, NP-marking, and cross-referencing. In addition, verbs sometimes 'register' the presence of an NP with a given grammatical function, without specifically identifying which NP has that function. Furthermore, two different techniques can function together as a strategy.

1.2.1   Order and arrangement   This technique is familiar from English. It is the order of NPs in (1) relative to the verb that indicates which is the subject (and therefore the agent) and which the object (and therefore the patient). English is an example of what we will call a 'fixed' word order system, one in which grammatical principles to a considerable extent prescribe the order of NPs. In such systems we find a 'basic' order, with various alternative orders systematically related to it. Since the workings of such systems are familiar from English, there is no need to discuss them here.

We also find systems in which there is a preferred order, but where a great deal of variation is possible as long as ambiguity is not introduced (although some languages seem to tolerate surprising amounts of ambiguity). Thus in Dakota (Van Valin (1985:366–7)), the preferred order is subject-object-verb

(SOV). If the semantics of the verb is not sufficient to determine which NP takes which role, this order is obligatory. Hence changing the order of the NPs in (9) changes the meaning:

(9)　　a. Wičása ki　mathó wą ktę
　　　　　man　　the bear　a　killed
　　　　　'The man killed a bear'

　　　　b. Mathó wą wičása ki　ktę
　　　　　bear　a　man　　the killed
　　　　　'A bear killed the man'

But if there is only one semantically plausible choice for subject, the relative order of NPs becomes free (though NPs and other constituents must remain in front of the verb):

(10)　　a. Wičása ki　ixʔé wą wąyále
　　　　　man　　the rock a　saw

　　　　b. Ixʔé wą wičása ki　wąyále
　　　　　rock a　man　　the saw
　　　　　'The man saw a rock'

In Dakota syntax, it does not seem to be sensible to try to describe the order possibilities in terms of a basic order and specific alternatives. Rather the order is flexible, subject to an SOV preference, especially when needed to prevent ambiguity. This sort of system we will call 'fluid', as opposed to the highly determinate word order system of languages like English.

　　Fluidity seems to be characteristic of many languages of diverse word order types. Fluid word order is usually not actually free, but is rather signalling pragmatic functions rather than grammatical relations. See Kiss (1987) and King (1995) for recent studies of two such 'discourse-configurational' languages, Kiss (1995) for a collection of studies, and Choi (1999) for detailed analyses of the phenomenon in Korean and German. The main difficulty in assessing the fluidity of word order is the fact that elicitation of sentences from informants will tend to produce the normal word order rather than a full spectrum of possible variants. Observation of actual language use, and examination of narrative and other natural genres of texts, will often reveal a much wider range of orders in their appropriate contexts.

1.2.2　NP-marking　　No language makes exclusive use of ordering to code grammatical relations, and many make very little use of it for this purpose. A technique which every language uses to some extent, and some use almost exclusively, is NP-marking. In this technique, the syntactic function of an NP is

indicated by a morphological marker on the NP. This marker may take the form of an inflection (see Bickel and Nichols, vol. III, chapter 3), or be a morphologically autonomous element, such as a clitic (which might also be called a 'particle'), a preposition (if it precedes the NP), or a postposition (if it follows). Both the inflections and the morphologically autonomous elements are often called 'case-markers'.

There is a great deal of fluctuation in the literature as to whether morphologically autonomous NP-markers are called 'particles', 'pre- or postpositions', or 'case-markers'. But there is widespread agreement that they should be seen as instances of a general technique which Nichols (1986) calls 'dependent marking', where the existence of a grammatical relation between two elements of a sentence is indicated by a marker placed on the dependent term. Dependent-marking can, however, apply to more than just NPs, for example to clauses or predicate adjectives.

In English the principal use of NP-marking is with prepositional phrase arguments and adjuncts. Thus the sentences of (11) are virtual paraphrases:

(11)   a. Bobby spoke to the meeting about the proposal
        b. Bobby spoke about the proposal to the meeting

To marks its NP as the addressee of *speak*, *about* marks its NP as the subject matter of the talk. Although the former order is preferred, both are possible, and it is clear that order does not mark the roles of these NPs.

Many languages make far more extensive use of NP-marking, using it to mark almost all NP functions, including subject–object or their counterparts. One example of this is Tagalog, which will be discussed below. Here we shall discuss an even more extreme example, Warlpiri, a Pama-Nyungan language of Central Australia (Hale (1973); Simpson (1991)).

In English, principles of order and arrangement not only indicate the functions of NPs, but the NPs themselves are also identified by means of such principles, since their constituent parts appear in a definite order, which can be described by phrase-structure rules, as explained in any reasonable introduction to generative grammar. In Warlpiri, both the functions and the constituency of NPs are usually indicated by NP-marking.

The one major principle of word order for Warlpiri simple clauses involves the 'auxiliary element'. This expresses the verbal categories of tense and mood (and also carries person–number markers for some of the verbal arguments, as we shall see in the next subsection), and comes in first or second position, depending on its phonological shape (Hale (1973:311–14); Simpson (1991:65)). The order of all other elements is free. Furthermore, there is no requirement that the constituents of an NP be contiguous; they must merely share the same endings.

The following three strings are therefore fully synonymous, and may be regarded as three versions of the same sentence:[6]

(12) a. Kurdu-ngku ka   maliki   wita-ngku wajili-pi-nyi
         child-ERG   PRES dog(ABS) small-ERG running-attack-NONPAST

      b. Wajili-pi-nyi ka wita-ngku maliki kurdu-nkgu

      c. Maliki ka kurdu-ngku wajili-pi-nyi wita-ngku
         'The small child is chasing the dog'

The auxiliary *ka* indicates that the tense is present. It is supplemented by the tense-ending on the verb, which shows nonpast tense. The ergative ending *-ngku* on *wita* 'small' and *kurdu* 'child' marks these as comprising one NP that bears A function. The absence of any ending on *maliki* shows that this belongs to a different NP, which can bear P function (we will see below that the absence of marking is also a characteristic of S function). This unmarked form is called the 'absolutive'. The endings thus indicate how the NP components are to be grouped together, and what function the resulting NPs are to have. There are twenty-one more arrangements of the words of (12), with the auxiliary in second position, and they are all grammatical and mean the same thing as (12).

There are two further observations to be made. First, *-ngku* is not a subject marker, because it is not normally used for NPs in S function. Rather, single arguments of one-argument verbs are normally in the absolutive case, with no marker:

(13)   Ngarrka   ka   purla-mi
        man(ABS) PRES shout-NONPAST
        'The man is shouting'

If we assume that the case marking directly reflects grammatical relations, we would have to deny that Warlpiri had a subject relation: rather, we would have to say that it had one grammatical relation covering A function, and another covering P and S functions. In fact, although they are not directly marked by the case forms, Warlpiri does seem to have subject and object grammatical functions, as we shall see in 3.1.4 below.

The second observation is that Warlpiri can group the members of an NP into a single overt constituent, and in this case the ending need only appear on the last word of the NP:

---

[6] Warlpiri, like many languages, lacks systematic indication of definiteness. The articles in the translations are arbitrarily chosen as 'the'. This will also be the case in the treatment of other languages, unless there is specific indication that definiteness is relevant.

(14)   a. Wita kurdu-ngku ka   maliki   wajili-pi-nyi
          small child-ERG   PRES dog(ABS) running-attack-NONPAST
          'The small child is chasing the dog'

       b. Wita   ka   kurdu-ngku maliki   wajili-pi-nyi
          small(ABS) PRES child-ERG dog(ABS) running-attack-NONPAST
          'The child is chasing the small dog'

The position of *ka* after *wita kurdu-ngku* in (14a) indicates that these two words form a constituent, and that they are therefore taken together as an NP despite the difference in endings. In (14b), where *ka* appears between *wita* and *kurdu-ngku*, these two words do not form a constituent, so *wita* has to be construed with *maliki*, and the sentence means 'the child is chasing the small dog'.

Warlpiri requires a somewhat more abstract kind of analysis than that which we have so far required for English: English NPs can be identified as units in a 'surface constituent structure' directly reflected in the linear order of elements. In Warlpiri we need at least two levels of analysis: overt constituent structure, relevant for auxiliary placement and a few other things, and a deeper level at which 'functional' units such as NPs are recognized even if their constituent elements are scattered throughout the overt structure.

1.2.3   Cross-referencing   In cross-referencing, also called 'agreement', various grammatical properties of an NP, such as noun-class (gender), number, person or case are registered on a word bearing some specific syntactic relation to the NP. As mentioned above, the Warlpiri auxiliary cross-references certain grammatical functions by hosting markers for their person and number. Third person singular ergative and absolutive NPs take no marker, so overt cross-referencing does not appear in examples (12–14). But first or second person, and dual or plural NPs, take non-null markers, as illustrated in example (15):

(15)
Nya-nyi   ka-rna-palangu   wawirri-jarra   (ngajulu-rlu)
see-NONPAST PRES-1SG(SUBJ)-3DU(OBJ) kangaroo-DU(ABS) (1SG-ERG)
'I see two kangaroos'

The clitic *rna* is here cross-referencing a first person singular A, *palangu* a third person dual P. In fact, as we shall see in 3.1.4 below, *rna* would also be used to cross reference an S, while a different clitic, *-ju* would be used for P, so the Warlpiri cross-referencing system is sensitive to subjects and objects, and provides some of the evidence that these are present, in spite of the case marking.

In contrast to case marking, where the marker appears on the dependent element, in cross-referencing it appears on the head, so this technique was

classified by Nichols (1986) as a kind of head-marking. Head-marking in Warlpiri and most other languages doesn't function primarily to code the grammatical function of NPs. In (15), for example, the markers are redundant because the functions are already coded by the markers on the NPs themselves (dependent-marking). Furthermore, in examples such as (12–14), where A and P are both third person singular, the markers are both zero, and thus provide no information at all about the functions of the NPs. Furthermore, in many languages, it is the case that most clauses have no overt NPs, so the cross-reference markers cannot be indicating their function. Rather the primary function of cross-referencing is to perform the function of pronouns.[7] Thus, in (15), the A pronoun *ngajulu-rlu* '1SG-ERG' is optional, and the meaning doesn't change if it is omitted. The P *wawirri-jarra* 'two kangaroos' is also optional, but if it is omitted the sentence means 'I saw them two'. A sentence such as *nya-nyi ka-rna* would mean 'I saw him/her/it': the absence of any cross-reference markers for P indicates that the P is third person singular.

Thus cross-referencing in Warlpiri (and most other languages that have it) is not a major part of the system for coding the syntactic functions of overt NPs. But since cross-reference markers often serve as substitutes for NPs, they are an important part of the system which specifies what entities take what roles in the situation denoted by the predicate. Since grammatical functions of NPs and the devices coding them are also part of this system, cross-referencing systems need to be investigated together with the more central NP function coding systems.

Occasionally, however, cross-referencing does provide the sole overt cue for the grammatical relation of an overt NP in a sentence. A good example is provided by Ancient Greek. Ancient Greek had case marking and very free word order (at least in writing). There is a participial construction in which the subject of the complement is suppressed when it is identical to some NP in the main clause. But the information is not lost as to what the subordinate clause subject is, because the participial verb form that the construction uses is marked for the gender (see below), number and case of the matrix NP that is to be understood as its subject. This information, especially the case information, is usually sufficient to identify what is to be understood as the subject of the complement.

It is thus the cross-referencing on the participle that disambiguates the following pair of sentences, by indicating the case of the S of the participle. Gender and number are also indicated, but these are the same (masculine singular) for both of the potential SS for the participle. Only the case is glossed, since the NPs that might be the S of the participle have the same gender (masculine) and number (singular):

---

[7] See Givón (1984a:353–85) for discussion of the close connections between pronominalization and cross-referencing, which Givón claims are in fact the same thing.

(16)   Klearchos   ape:nte:se Philippo:i   apio:n
        Klearchus(NOM)   met   Philip(DAT) leaving(NOM)
        'Klearchus met Philip while Klearchus was leaving'

(17)   Klearchos   ape:nte:se Philippo:i   apionti
        Klearchus(NOM)   met   Philip(DAT) leaving(DAT)
        'Klearchus met Philip while Philip was leaving'

This is an unusually straightforward example of cross-referencing marking grammatical relations. Usually, when cross-referencing manages to do this, it does so by means of complex interactions with other techniques and principles.

A particularly complex and interesting case of this are the 'obviation and inverse-marking' systems originally found in Algonquian languages, and then more widely.[8] The basic idea of these systems is that there are two third person categories, 'proximate' and 'obviative', where 'proximate' applies to an NP, unique at any particular point in the discourse, which is seen as the prime focus of attention (such as the protagonist of the current action), while 'obviative' applies to the other third person NPs. A normal 'direct' transitive verb with a third person subject and object then describes the proximate as acting on the obviative, while if the obviative is acting on the proximate, a specially marked 'inverse' form is used.

In Plains Cree, for example (Wolfart (1973); Dahlstrom (1991)), obviative NPs bear a marker *-ah* (it is clear that this does not mark case or grammatical function, but a kind of discourse status), while proximates are unmarked. In (a) below, the obviative is the patient, and the verb is 'normal' (non-inverse), whereas in (b) the agent is obviative, and the verb is inverse in form:

(18)   a. aya·hciyiniw-ah nisto e·=mipah-a·t awa na·pe·sis
            Blackfoot-OBV   three kill-DIRECT   this boy
            'This boy had killed three Blackfoot'
                            Bloomfield (1934:98), cited in Dahlstrom (1991:62)

        b. osa·m   e·=sa·kih-ikot ohta·wiy-ah   aw o·skini·kiw
            too much   love-INV   his father-OBV this young man
            'for his father too much cherished this young man'
                            Bloomfield (1934:58), cited in Dahlstrom (1991:63)

Dahlstrom shows that the obviative marking on the nouns, and the direct/inverse marking on the verbs, is irrelevant to grammatical relations, the A being a subject and the P an object regardless of these markings. These systems also constitute a case of PTVs having two different-looking treatments of A and P, depending on which is the proximate in the discourse.

---

[8] See Aissen (1997, 1999) for discussion, and an application to the Mayan language Tzotzil, where obviation had not previously been seen as relevant.

So we have a combination of dependent-marking (obviation on the nouns) and head-marking (direct/inverse marking on the verbs) conveying the semantic roles. Cross-referencing also enters the mix: when a verb has first or second person arguments, these are cross-referenced in fixed positions on the verb, with the direct/inverse marking indicating which is A and which P:

(19)   a. ki-wa·pam-i-n
             2-see-DIRECT(1)-SG
             'You(sg) see me'

         b. ki-wa·pam-iti-n
             2-see-INV(1)-SG
             'I see you(sg)'   Dahlstrom (1991:42)

In this language, second person is treated as proximate as opposed to first, but the opposite ranking is also possible. The entire system comprises one kind of dependent-marking and two kinds of head-marking (cross-referencing together with direct/inverse marking), which all work together in a complicated way to signal the semantic roles.

1.3   *Pragmatic functions*

Pragmatic functions involve a great variety of considerations, many of which are not very well understood. Some of the important concepts are: (a) what the hearer is presumed to be already conscious of ('given' vs 'non-given'); (b) what the sentence is about ('topicality'); (c) whether an NP has or doesn't have a referent uniquely identifiable to the hearer ('definiteness' and 'identifiability'); (d) whether the speaker is referring to a particular instance of an entity as opposed to any instance of it ('specificity'); (e) what is 'foregrounded' as important vs what is 'backgrounded' as secondary; (f) the point of view taken by the speaker on the situation being talked about ('empathy', or 'perspective'); (g) inherent 'salience properties' of NPs, such as animacy, humanness, or first-personhood.

   Many of these concepts are discussed and clarified in Lambrecht (1994), and their interactions with sentence structure are examined in Foley in chapter 7. In this section we will limit ourselves to discussing three major 'pragmatic articulations' of sentence-structure that tend to have significance for grammatical functions: 'topic–comment', 'presupposition–focus' and 'thetic'. Pragmatic functions are relevant to grammatical functions because there are frequently rules or tendencies relating the two. 'Subjects', for example, as we will discuss later, often show either a strong tendency or even an absolute requirement to be topics (Lambrecht (1994:131–7)).

1.3.1   Topics and topic–comment articulation   Topics are generally thought of as entities previously known to the hearer, which it is the function of the sentence to provide some further information about (unfamiliar entities can, however, be introduced into the discourse and then become topics; this is the function of the thetic articulation, especially its presentational subtype). A sentence that has one or more topic entities can be said to have 'topic–comment' articulation. There are two principal kinds of topics: those whose topicality is predictable from the immediately preceding discourse, and those whose topicality is not. For an illustration of the two types, consider the following story:

Once upon a time there was a king with two sons. The older son expected to take over the kingship. *He* spent his time travelling with the army and working with the secret police. As for *the younger*, he concentrated on studying philosophy at the University.

 The italicized pronoun *he* in the third sentence is expected to be topic, since its referent is also the topic of the immediately preceding sentence. *The younger* in the fourth sentence represents a new, unexpected topic. The switch in topic is registered by the *as for* construction, which seems to indicate that some entity, introduced previously in the discourse, but not referred to recently, is being made the new topic. We might call these two types 'expected topic' and 'switch topic'. In many languages the subject grammatical relation is associated with the topic (expected or switch) function. This association can manifest itself as a requirement that subjects be definite – as discussed by Keenan (1976a:252–3), for Malagasy and Kinyarwanda, and Givón (1979:26–7) more generally – or as a tendency for them to be definite (Givón (1979:26–8)).

On the other hand, so-called 'topicalization' constructions are frequently (but not always) associated with switch-topic functions, as illustrated by the *as for* construction above.

We need to distinguish between a topic entity (the older or younger of the king's sons in the passage above, depending on what sentence is being analysed), and a topic expression (NP), such as *he* or *(as for) the younger* (Lambrecht (1994:127–8)). Expected topic entities tend to be expressed by reduced linguistic constituents, such as pronouns, or by nothing at all (this is called 'null anaphora'). Therefore, in some languages, it is common for sentences with a topic entity to have no topic expression, so that if we want to talk about a 'sentence without a topic', we need to be sure whether we're talking about topic entities or topic expressions.

The topic expressions then are the linguistic materials referring to the entities that the sentence is about; the comment is the remainder – that is, what the sentence actually says about them. If there is no topic expression, but there is a topic entity, then the entire sentence will constitute the comment expression.

1.3.2   Focus–presupposition articulation   In this kind of articulation, there are again two components. One, the presupposition, presents incomplete information about a situation of which the speaker presumes the hearer to be aware. The other, the focus, is the missing information, which the speaker presumes that the hearer wants to know. The so-called *it*-cleft construction of (2), repeated below for convenience, is a typical example of focus–presupposition articulation:

(2)   a. It is the farmer that kills the duckling
        b. It is the duckling that the farmer kills

As was pointed out at the beginning of the chapter, in (2a) 'the farmer' is the focus, and 'kills the duckling' is the presupposition. The speaker assumes the hearer knows that someone or something killed the duckling, and gives the information that it was the farmer that did it.

English has two other extensively discussed focus–presupposition structures, the *wh*-cleft construction and 'contrastive stress' on the focus:

(20)   a. A bear is what the man killed
        b. The man killed *a bear*

In (20a), 'a bear' is the focus, and 'what the man killed' is the presupposition. The speaker assumes that the hearer knows that the man killed something and tells the hearer that this was a bear.

All three constructions differ in their usage. To see a difference between either kind of clefting and contrastive stress, observe that (20b) is a better answer to the question *What did the man kill?* than either (20a) or its *it*-cleft counterpart *It's a bear that the man killed.* See Prince (1978) for the differences between the two cleft constructions.

Topic–comment articulation can be superposed on focus–presupposition articulation: in a sentence such as *George is looking for* BEARS, *George* might be the topic (so the NP *George* would be a topic expression), and *bears* an expression of the focus. The comment is expressed by *is looking for bears*, the presupposition by *George is looking for X*. Some languages such as Tzotzil, allow both to be marked simultaneously (Robinson (2002)).

1.3.3   Thetic articulation   Not all sentences have topic–comment or presupposition–focus articulation. A less studied third alternative, recently emphasized by Lambrecht, is 'thetic articulation'. In thetic articulation, the entire sentence can be taken as a comment whose topic is the ambient situation rather than some specific, delineated component thereof that has been accepted as something to talk about. Lambrecht illustrates thetic articulation with the

contrast between (21a) and (21b), emphasis represented by small capitals:

(21)   a. my car BROKE DOWN
       b. MY CAR broke down

Example (21a) could be used to answer a question such as *Where is your car?*, which would establish the car as a suitable topic to deliver more information about. Example (21b) on the other hand would be quite inappropriate for this purpose. What it would be good for is presenting as an excuse upon rushing into a meeting twenty minutes late, where (a) would on the other hand be out of place. In such a case the car is not the topic, but part of the comment, an explanation of the present situation, which is the actual topic.

In English, thetic subjects receive stress relative to the verb phrase, but in some languages, such as French, they are just impossible. In French, it seems to be the case that subjects must be topics. Hence in (22a), the French counterpart of (21a), the car, which is topical, is the subject just as it is in English, while in (22b), the French counterpart to (21b), the car, which is thetic rather than topical, must be expressed as an object:

(22)   a. Ma voiture est EN PANNE
          My car   is   broken down
          'My car BROKE DOWN'

       b. J'ai   MA VOITURE qui   est en panne
          I. have my car   which is broken down
          'MY CAR broke down'

A more widely discussed subtype of thetic articulation is presentational articulation, used to announce the existence or appearance on the scene of a hitherto unknown entity:

(23)   a. There's a snake in the shower
       b. Once upon a time there was a king with three children

Although English has the special presentational construction illustrated above, it is also possible for presentational subjects to appear with no special marking (other than thetic stress):

(24)   a. A king with three children lived in a valley
       b. A person is standing outside the door

But languages with a restriction that subjects be topics always need to use a special construction for sentences with thetic articulation, as illustrated above by French in (22).

2   Overview of grammatical functions

With these preliminaries completed, we will proceed to look at the grammatical functions themselves. We will first present a general classification of the types of grammatical function, then examine specific types in greater detail. Figure 3.2 is a diagram of the taxonomy of grammatical functions that we will be looking at.[9]
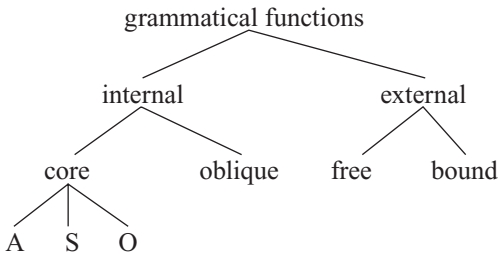
Figure 3.2 Taxonomy of grammatical functions

2.1   *Types of grammatical function*

We will distinguish three fundamental types of grammatical function, core, oblique and external, which may be thought of as constituting successive layers of clause structure. The first division is between the external functions and the others, which we will call internal.

External functions give the appearance of being essentially outside of the basic clause structure, and are each associated with a fairly specific pragmatic function. The *it*-cleft construction of (2) and the *as for* construction above illustrate typical external functions. An external function never itself has an association with any specific semantic role, although the NPs bearing them often (but not always) acquire a semantic role by other means.

The internal functions have close associations with semantic roles, though they may be associated with pragmatic functions as well. Subject, object and the various prepositional phrases in (7) and (8) bear typical internal functions. Note that by saying that internal functions are associated with semantic roles we do not mean that they have them as invariant properties, but merely that they tend to go together. Subject in English is associated with the semantic role of agent, but many subjects are not agents; the preposition *to* is often associated with the semantic role recipient, but not always.

Among internal functions, A, S, and P have a special status, because they almost always have a variety of properties which set them off from most of

---

[9] I'm indebted to Stuart Robinson for suggesting and providing this diagram.

the other grammatical functions. In English for example, with the exception of personal pronouns, A, S, and P are unmarked NPs, with functions coded by order relative to the verb, while most other functions are coded by prepositional NP-marking.

In English, not only do NPs with A, S and P functions differ in appearance from prepositionally marked NPs, they also differ in various aspects of their syntactic and semantic behaviour. Two especially important properties are that they tend to express a wider range of semantic roles, and that they tend to be 'targetted', that is, singled out for special treatment, by various rules of syntax which appear to function in terms of specific grammatical relations, rather than in terms of semantic roles or pragmatic functions. For example, subjects are omitted in various kinds of nonfinite subordinate clause constructions, such as the infinitive complement of *want* in (25a), and the participial adjunct in (25b):

(25)   a. John wants to buy a new computer
       b. Having bought a new computer, John couldn't afford lunch for
           three months

On the other hand objects may be passivized:

(26)   a. John was arrested
       b. John was given a book

Rules involving prepositional phrases (hereafter in this chapter abbreviated to 'PP'), on the other hand, tend to apply to a wide range of constituents, including non-PPs, with restrictions being statable in terms of semantically specifiable categories rather than syntactic ones.

In most other languages there is a similar distinction between a small class of grammatical relations expressing A, S and P (and sometimes other) functions, which behave somewhat like subject and object in English, and a larger class, which behave like English PPs. We thus divide the internal function into two categories, calling the former class of grammatical functions 'core', the latter, 'oblique'. Thus the core functions are by definition A, S, P, and whatever other grammatical functions are sufficiently like them to be plausibly grouped with them and opposed to the others, which are the oblique functions.

Languages in which the core/oblique distinction corresponds to that between bare NPs and those carrying a marker are not uncommon. Some additional examples are Jacaltec and other Mayan languages (Craig (1977); England (1983a)), Bahasa Indonesia (Chung (1976)), Dakota (Van Valin (1985)), and the Bantu languages, some of which will be discussed below. In other languages, there does not seem to be a significant syntactic distinction between marked and unmarked NPs. In Japanese (Kuno (1973)), Russian (Comrie (1979)), and Tagalog (Schachter and Otanes (1972)), for example, all NPs are marked. In other languages, such as Warlpiri, some NPs are unmarked, but the marked NPs

include some which are by definition core (A in Warlpiri). Furthermore, there is no striking overall difference in syntactic behaviour between the marked and the unmarked NPs.

Nonetheless, something corresponding to the core/oblique distinction in English usually seems to exist even in languages where A, S and P normally carry the same kinds of markers as other grammatical functions. One set of cases, commonly called 'syntactic', 'structural', or 'direct' cases, mark the core functions, another, commonly called 'semantic' cases, mark the oblique functions. NPs with syntactic cases tend to express a wide range of semantic functions and to be targetted by rules sensitive to grammatical function, while NPs with 'semantic' cases tend not to have these properties.

Usually, the properties of core NPs suggest that they should be viewed as bearing 'abstract grammatical relations': structural relationships which are not necessarily directly reflected by coding features, and do not necessarily correlate precisely with semantic roles, pragmatic functions, or other aspects of meaning. By contrast, the grammatical function of obliques, such as the PPs in (7–8) can for the most part be identified with their semantic roles.

Most of the typological work on grammatical functions has been directed toward core functions, although recently there has been increasing consideration of external ones. Obliques on the other hand still seem to be relatively neglected. In the remainder of this section we will briefly consider external and oblique functions, and then, in section 3, we will discuss at greater length core functions and the grammatical relations associated with them.

2.2   *External functions*

As we observed above, external functions give the appearance of being essentially outside of the clause structure, and are each closely associated with a specific pragmatic function. But the grammar of a language does not specify any associations between external functions and semantic roles (ways of participating in the situation described by the sentence), and, for some external functions, their bearer needn't have any semantic role in the sentence at all.

Suppose Jim's wife, Harriet, has left him. If some of the couple's former friends were discussing Jim, one of them might say:

(27)   Speaking of Jim, what's Harriet been up to lately?

Jim is brought up as the topic of the sentence, what the sentence is about, but does not have a semantic role with respect to the predicate. In English, such constructions have a fairly minor place in the system of the language, but in many languages they are the predominant form of sentence in ordinary usage. Such languages were called 'Topic Prominent' by C. N. Li and Thompson (1976), and seem to be especially characteristic of Southeast Asia.

We illustrate typical instances of such constructions with examples from Chinese, Lahu (Tibeto-Burman), and Japanese, with the topic (which appears in initial position) italicized:

(28)   a.   Chinese
        Neì-chang huɔ xìngkui   xīaofang-duì laí   de   kuài
        that-CLSFR fire fortunate fire-brigade   come PTCL quick
        'That fire, fortunately the fire brigade came quickly'
                                        C. N. Li and Thompson (1976:482)

       b.   Lahu
        Hɔ   ɔ̄   na-qɔ́ yï   ve   yò
        elephant TOP nose long PICL DEC
        'Elephant, noses are long'
                                        C. N. Li and Thompson (1976:482)

       c.   Japanese
        Nihon   wa Tokyo ga   sumi-yoi
        Japan   TOP Tokyo NOM easy.to.live.in
        'As for Japan, Tokyo is comfortable to live in'   Kuno (1973:65)

These examples cannot be adequately glossed in English, since their nearest counterparts use constructions using *as for* and *speaking of*, which, as noted above, carry a switch-topic force that is absent in the examples of (28). Chafe (1976:50) characterizes the function of the topic in these constructions as that of setting 'a spatial, temporal or individual framework within which the main predication holds' (see also Lambrecht (1994:118)).

External functions whose bearers needn't have a semantic role in the accompanying clause will be called 'free'. Free external functions always seem to introduce topics, functioning more or less as described by Chafe. Furthermore they always place an NP at the beginning of the sentence, either with accompanying morphological material (Lahu, Japanese, English) or without it (Chinese).

Other external functions require their bearer to have a semantic role in the clause (of course this is also possible for free topics). We call these 'bound'. In English the *it*-cleft construction is a bound external function, as is the 'topicalization' construction in which an NP is preposed without additional marking. Observe the contrast below:

(29)   a.   As for American self-confidence, Columbia gave people a lift
       b.   *American self-confidence, Columbia gave people a lift
       c.   *It was American self-confidence that Columbia gave people a lift

In all of these exampIes. the clause fails to assign a semantic role to the initial NP. The result is acceptable in the case of the *as for* construction (29a), but not in the case of the others, the topicalization construction (29b) and the *it*-cleft

construction (29c). This illustrates that the *as for* construction is a free external function, while the topicalization and *it*-cleft constructions are bound external functions.

Bound external functions have a wider range of pragmatic effects than free ones, such as marking focus and presupposition or presentational articulation, and they are coded by a wider range of techniques, including movement to various positions in the sentence-structure, and also marking *in situ*, without any special position. This latter possibility is illustrated below for the Dravidian language Malayalam.

Malayalam (Mohanan (1982)) is an SOV language with NP-marking by means of case markers and postpositions, and therefore, as one would expect, has fairly free word order (but, unlike Warlpiri, major constituents such as NPs cannot be broken up). There is a 'cleft' construction in which the verb is suffixed with *at* 'it', and the clefted NP is suffixed with a form of *aa* 'be'. The normal word order for this construction is the same as in a non-clefted sentence. Below we give a sentence in normal word order, together with four clefted variants. the cleft NPs being italicized:

(30)   kuṭṭi   iṉṉale   ammakkə   aanaye   koṭuttu
        child(NOM)   yesterday   mother(DAT) elephant(ACC) gave
        'The child gave an elephant to the mother yesterday'

(31)   a. *kuṭṭiy-aanə*   iṉṉale   ammakkə   aanaye   koṭutt-atə
           child(NOM)-is   yesterday   mother(DAT) elephant(ACC) gave-it
           'It is the child that gave an elephant to the mother yesterday'

        b. kuṭṭi   *iṉṉaley-aanə* ammakkə   aanaye   koṭutt-atə
           child(NOM)   yesterday-is   mother(DAT) elephant(ACC) gave-it
           'It is yesterday that the child gave an elephant to the mother'

        c. kuṭṭi   iṉṉale   *ammakk-aanə* aanaye   koṭutt-atə
           child(NOM)   yesterday   mother(DAT)-is elephant(ACC) gave-it
           'It is the mother that the child gave an elephant to yesterday'

        d. kuṭṭi   iṉṉale   ammakkə   *aanayey-aanə* koṭutt-atə
           child(NOM)   yesterday   mother(DAT) elephant(ACC)-is gave-it
           'It is the elephant that the child gave to the mother yesterday'

It is also possible to cleft the verb, although this does not concern us here. The elements of all of these sentences could be freely reordered.

In addition to rearrangements and markings, external functions can lead to the appearance of a variety of further subtle effects in the clauses they occur in (Zaenen (1983)). Nonetheless it is clear that they are relatively independent of the system of internal grammatical relations that provide the primary expression of semantic roles, and are in effect 'superposed' on it.

Sentence-level intonational and stress features, operating either alone or in conjunction with syntactic mechanisms, may also be employed to express bound external functions. In English, for example, we can impose focus–presupposition articulation simply by stressing the focus:

(32)   The farmer kills the DUCKLING (c.f.(2))

Stress is frequently used as a focus marker. On the other hand it does not seem to be used to mark topics, except in contrastive constructions:

(33)   Speaking of Mary and Jim, MARY will like this dish, but JIM will hate it

This is presumably because topics are familiar information with relatively less need for attention to be directed to them, while foci are the new information that is actually being communicated.

2.3   Oblique functions

In this section we examine oblique grammatical functions. We will first investigate English, showing that English obliques fall into two main classes: arguments and adjuncts. The distribution of arguments is governed by potentially idiosyncratic specifications on verbs (or other predicates). Adjuncts on the other hand appear whenever they would be semantically appropriate. In fact, we shall see that it is reasonable to think of the argument/adjunct distinction as overlapping the core/oblique distinction, with all core NPs and some obliques being included in the class of arguments. Adjuncts, on the other hand, always seem to be oblique, in that they do not seem to exhibit behavioural similarities to A, s and P.

Then we will look at obliques in Warlpiri, to illustrate something of the behaviour of obliques in a case-marking language. Finally we will briefly summarize the dimensions of typological variation in systems of oblique grammatical functions.

2.3.1   Obliques (PPs) in English   English oblique NPs are usually expressed within prepositional phrases (PPs), except for certain time expressions, where a preposition does not have to be expressed: *Mary left the next day*. English PPs are not homogeneous but seem to fall into classes, which can be defined in terms of the way in which their form and distribution is or is not determined by the verb. As stated above, the two principal classes are what we shall call 'arguments' and 'adjuncts'. The distribution of adjuncts is not subject to idiosyncratic restrictions imposed by the predicate, but only to the requirement that the sentence make sense. The circumstantial roles of section 1.1.2 are often introduced by adjuncts. Thus in English, any verb which is semantically suitable may take a locative

phrase, or a benefactive phrase with the preposition *for*. For example, the reason that example (34b) is odd is not because of some syntactic restriction on adjuncts expressing reasons, but rather because tree branches don't have minds, and therefore lack motives for doing things:[10]

(34) a. John prodded the snake for fun.
      b. # The branch fell off the tree for fun.

In contrast, the distribution of arguments is subject to idiosyncratic restrictions imposed by verbs. To see the nature of these restrictions, let us examine the nature of the constructions associated with verbs of giving, such as *give*, *hand*, *present*, etc., in which an agent transfers a theme from his/her own custody to that of a recipient.

Such verbs take six patterns of association between their semantic roles and the grammatical relations that express them, as illustrated below. The patterns in (a) and (b) are the major ones, (c) is minor, and (d–f) are extremely minor:

(35) a. Susan handed Paul the shovel
      b. Susan handed the shovel to Paul
      c. They supply us with weapons
      d. Cheech laid a joint on Chong
      e. Geraldine foisted six kittens off on(to) Jock
      f. J. R. bestowed many favors (up)on Afton

Examples (35b–f) illustrate various oblique constructions, while (35a), with two bare NPs after the verb, illustrates something we haven't discussed yet, a 'double object' or 'ditransitive' construction. In section 3.2.1 below we will argue that the first postverbal NP is a 'primary object' bearing the same grammatical relation as the sole object of a transitive verb, while the second bears another core grammatical relation, 'secondary object'. See section 2.3 of chapter 4 by Dryer, for further discussion of ditransitives.

There is considerable systematicity in the relations between semantic roles and their overt expressions in (35). In the double-object and *with* constructions, recipients are the first or sole objects. Otherwise they are the objects of goal prepositions such as *to*, *on*, *into* and *onto* (the latter three sometimes being optional alternants). Themes, on the other hand, are primary objects (35b, d, e and f), second objects (35a), or objects of *with* (35c).

But there is also considerable idiosyncrasy. *Hand*, and a great many other verbs, appear in patterns (a) and (b), but not the others. *Supply* appears with (c) and (b), and maybe (a) for some speakers, but not with (d–f). *Equip*, on the other hand, appears only in (c). None of the verbs taking any of (a–c) take

---

[10] The '#' mark in front of the (b) example indicates that the example is semantically bizarre, but not ungrammatical.

any of (d–e), except *fob off*, which takes (c) and (d), with substantially different meanings: *Fred fobbed Jack off with a scratched CD* vs *Fred fobbed a scratched CD off on(to) Jack* (in the first sentence, Fred is getting rid of Jack, in the second, a CD).

There are, furthermore, idiosyncratic restrictions on whether some of these obliques are optional or obligatory. The *with* phrase can be ellipsed with *supply*, but not *provide*, with the object retaining the receipient role:

(36)   a. We supply Iran (with weapons)
          b. We provide Iran *(with weapons)

Similarly, *to*-objects are usually optional, but with some verbs they are obligatory:

(37)   a. Susan passed the shovel (to Paul)
          b. Susan handed the shovel *(to Paul)

There has been substantial recent work, such as Pinker (1989) and Wechsler (1995), on how to predict the choice of preposition, and whether the PP is obligatory or optional. But some facts of preposition choice seem to resist explanation (Wechsler (1995:122)), as do some of the optionality facts, such as those in (36) and (37) above. So there still seems to be a category of obliques that are subject to lexical control, and which therefore may be reasonably regarded as a kind of argument.

Furthermore, even in the great majority of cases, where the choice of preposition is semantically predictable, we can make a case that it is not making an independent contribution to the meaning, since one cannot vary the choice of preposition independently to vary the meaning. This suggests that the verb is in some sense determining the semantic role of the NP, which is in addition being marked by the preposition. Such a view is indeed taken by Wechsler (1995), following earlier work by Gawron (1986) and Jackendoff (1990).

But there are also PPs which appear to be arguments where the preposition does seem to make an independent contribution to meaning. The verb *put*, for example, takes an obligatory directional phrase in *in(to)* or *on(to)* (and most other goal PPs), while *move* takes an optional directional PP in *into* or *onto*, but not *in* or *on*:

(38)   a. Cally put the key *(on(to) the table / in(to) the box)
          b. Cally moved (the computer) (on*(to) the table / in*(to) the box)

The *in/on* components here indicate spatial relationships, while the possibilities for omitting or including *to* seem more arbitrary. (*On/in* is of course acceptable with *move* when the PP is an outer locative rather than a directional.)

These PPs seem clearly to be arguments rather than adjuncts, but they resemble adjuncts in that the preposition is a partially independent bearer of meaning.

It seems appropriate to think of the PP as a whole as being an argument to the verb, rather than of the NP within it as being the argument, with the preposition marking its role.

We therefore classify English PPs into adjuncts, and two types of arguments. In the first kind of argument, which we will call 'P-objects', the verb determines the choice of preposition, and the NP within it functions as an argument of the verb. In the second type, which we will call 'P-complements', although the verb may constrain the choice of preposition, it does not determine it completely. Rather the preposition expresses meaning to some extent independently from the verb, and the PP as a whole functions as an argument.[11]

Although in many cases it is clear whether one is dealing with an argument or an adjunct, there are also doubtful (perhaps intermediate) cases. For example, almost any verb which is semantically appropriate may take an instrumental *with*-PP, which suggests that these are adjuncts:

(39)   a. The old man walks with a stick
        b. Marcia watched the koalas with binoculars
        c. Jimmy poked Owen with a stick

But P. H. Matthews (1981:18) notes that the verb *go* does not take instrumental *with*: *He went with a stick* means merely that he went carrying a stick with him, not that he used it as an instrument in the activity of going. It is not clear whether this restriction can be made to follow from the meanings of *go* and instrumental *with*. Therefore it is unclear whether instrumentals should be regarded as arguments or adjuncts.

Drawing the argument/adjunct distinction may require considerable knowledge of a language, and deep insight into its semantics. The core/oblique distinction, on the other hand, is usually relatively obvious, although in a few cases it too is somewhat obscure. For this reason the latter rather than the former distinction is emphasized in this study.

Oblique grammatical functions are typically more tightly tied to specific semantic roles than are the core grammatical relations. In the case of the adjuncts and P-complements, the NP-marker of the oblique grammatical function specifies the semantic role to a considerable degree independently of the verb, while the P-object markers are also more tightly tied to given semantic roles than are subject or object. Objects, for example, can be Themes or Recipients, while *with*-objects can be Themes but not Recipients, and *to*-objects can be Recipients but not Themes. To get a better sense of the nature of the core/oblique distinction, we will next examine obliques in Warlpiri, a language where the

---

[11] The terms 'P-object' and 'P-complement' are borrowed from Bresnan (1982); see Bresnan (2001:275–80)) for more recent discussion.

Table 3.1 *Warlpiri cases*

Local semantic		Non-local semantic	
Locative ('at'):	*-rla/-ngka*	Instrumental:	*-rlu/-ngku*
Allative ('to'):	*-kurra*	Causal:	*-jangka*
Elative ('from'):	*-ngurlu*	Considerative:	*-wana-wana*
Perlative ('along'):	*-wana*		
Comitative ('with'):	*-rlajinta*		
Derivational		Syntactic	
Associative	*-warnu*	Ergative:	*-rlu-ngku*
Excessive	*-panu*	Dative:	*-ku*
Denizen of	*-ngawurrpa*	Absolutive:	-∅
Like	*-piya*		
Possessive	*-kurlangu*		
Privative ('without')	*-wangu*		
Proprietive ('having')	*-kurlu, -pamta, -manji*		
Source	*-jangka*		

core/oblique distinction does not correspond to that between morphologically marked and unmarked NP.

2.3.2   Obliques in Warlpiri   Warlpiri cases (NP-markers) can be divided into two main groups: the 'syntactic' cases (ergative, dative and absolutive) and the 'semantic' cases (all the rest). The latter can be further divided into three subgroups: local semantic, non-local semantic, and 'derivational' semantic. The syntactic cases code core functions, which will be reviewed for Warlpiri in section 3, comprising all core NPs and some obliques. The local and non-local semantic cases express oblique functions, with the local semantic cases expressing primarily spatial notions, the non-local cases non-spatial ones (the local cases also have some non-spatial uses). The 'derivational' cases seem for the most part to form modifiers of NPs rather than arguments or adjuncts of the verb, and are therefore largely beyond the scope of this chapter.

Table 3.1 presents some of the most important cases. The listing for the non-local semantic cases is incomplete, since the boundary of this category is unclear. The endings *-ngka* (LOC) and *-ngku* (ERG/INSTR) are used after stems with two syllables, *-rla* (LOC) and *-rlu* (ERG/INSTR) after stems with three or more. Furthermore, the form of some of these endings is affected by a vowel harmony rule converting *u* to *i* after stems in *i*, so that we get *maliki-ki* 'dog-DAT', *wati-ngki* 'man-ERG', and *yuwarli-ngirli* 'from the house'.

The local semantic cases primarily indicate the spatial notions of location at (or on, or in), motion to, motion from, motion along, and motion or position

together with:

(40)   a. Lungkarda   ka   ngulya-ngka nguna-mi
           bluetongue(ABS) PRES burrow-LOC lie-NONPAST
           'The bluetongue skink is lying in the burrow' (locative)

        b. Nantuwu   ka   karru-kurra parnka-mi
           horse(ABS) PRES creek-ALL   run-NONPAST
           'The horse is running to the creek' (allative)

        c. Karli   ka   pirli-ngirli   wanti-mi
           boomerang(ABS) PRES stone-ELATIVE fall-NONPAST
           'The boomerang is falling from the stone' (elative/ablative)

        d. Pirli   ka-lu-jana   yurutu-wana   yirra-rni
           stone(ABS) PRES-they-them road-PERLATIVE put-NONPAST
           'They are putting stones along the road' (perlative)

        e. Maliki   ka   nantuwu-rlajinta   parnka-mi
           dog(ABS) PRES horse-COMITATIVE run-NONPAST
           'The dog is running along with the horse' (comitative)

Hale (1982) provides a detailed account of the semantics of these cases, Simpson (1991) a more formal analysis.

The Warlpiri case system makes fewer distinctions than the systems of prepositions of English, but similar effects are achieved by other means. There are, for example, adverbial particles which, although not syntactically bound to a local case-marked NP, nonetheless refine the locative concept expressed. *Kulkurru*, for example, specifies *between*-ness:

(41)   Maliki   ka   nguna-mi   yuwarli-jarra-rla kulkurru-jarra
        dog(ABS) PRES lie-NONPAST house-DU-LOC   between-DU
        'The dog is lying between the two houses'

Without *kulkurrujarra*, the sentence could be interpreted as meaning merely that the dog was near the houses.

Occasionally the local cases are used idiomatically, in ways not fully explicable in terms of their basic meanings. For example the verb *manyu-karri-mi* 'play-stand-NONPAST', meaning 'to play a game', takes the locative case on the game played. This may co-occur with a locative designating the place where the event happens:

(42)   Ngarrka-patu ka-lu   manyu-karri-mi   kardi-ngka karru-ngka
        man-PL(ABS) PRES-they play-stand-NONPAST card-LOC   creek-LOC
        'The men are playing cards in the creek'

These usages are reminiscent of idiomatic P-objects in English.

The non-local semantic cases are for the most part minor in the structure of the language. The 'true' instrumental expresses the instrument used by an agent to act on a patient. It *only* appears with transitive verbs taking an ergative agent and absolutive patient, not with intransitives (or with a category we shall discuss below of two-argument verbs not taking an ergative). See examples (43a) and (43b) below. There is another method for expressing the instrumental relation, and this one may be used with either transitives or intransitives. It involves one of the 'derivational' semantic cases, the proprietive *-kurlu* 'with'. The basic meaning of *-kurlu* is possession, but the meaning of it can be extended to indicate not only possession but use, as shown in examples (43c) and (43d):

(43)   a. Wawirri   kapi-rna kurlarta-rlu panti-rni   ngajulu-rlu
           kangaroo(ABS) FUT-1SG spear-INSTR spear-NONPAST 1SG-ERG
           'I will spear the kangaroo with a spear'

        b. *Purlka   ka   watiya-rlu warru-wapa-mi
           old man(ABS) PRES stick-INSTR around-walk-NONPAST
           'The old man is walking around with a stick'

        c. Ngarrka-ngku ka warlu   paka-rni   warlkurru-kurlu-rlu
           man-ERG   PRES firewood chop-NONPAST axe-with-ERG
           'The man is chopping firewood with an axe'

        d. Purlka   ka   watiya-kurlu warru-wapa-mi
           old man(ABS) PRES stick-with   around-walk-NONPAST
           'The old man is walking around with a stick'

Example (43c) could be interpreted as possessive for *-kurlu* rather than that of use, to give 'The man with an axe is chopping firewood' (using some other instrument), and so also (43d) 'The old man with a stick is walking around'. The instrumental sense is nevertheless the usual one in sentences such as these, expressing an action where the use of the object in question is in fact likely.

The ending listed as causal is also widely used to indicate source of motion (elative), and preferred as such by some speakers. But it also indicates the cause for the situation designated by the sentence, or a potentially causal prior event:

(44)   Ngarrka-patu ka-lu   warrki-jangka mata nguna-mi-lki
        man-PL(ABS) PRES-they work-CAUSAL tired lie-NONPAST-now
        'The men are lying down tired now after work'

It can also indicate the material out of which something is made: for example, from wood in 'They are making boomerangs from wood'. The 'considerative' (CONS) is applied to an NP denoting something that is given in exchange for something else:

(45)   Japanangka-rlu   ka-ju   karli   yi-nyi
         Japanangka-ERG   PRES-1SG(OBJ) boomerang(ABS) give-NONPAST
         miyi-wanawana
         food-CONS
         'Japanangka is giving me a boomerang in exchange for food'

This illustrates nicely that a serious account of semantic roles must go consid-
erably beyond the simple agent, patient, source, goal, etc., categories that were
defined in the introduction to this chapter above.

   The uses of the cases we have considered so far mostly involve qualifications
of some facet of the action of the verb: the path taken by some participant (and
thereby, in some sense, of the 'action'), or additional participant. They thus
express participatory semantic roles (see 1.1.2), and function analogously to
the oblique argument (P-objects and P-complements) of English. The exception
is the causal use of *-jangka*, which provides the background for the event, and
is thus a circumstantial adjunct.

   The principal circumstantial case is the locative, which can place an event in
space (already illustrated in (42) above) or in time:

(46)   Ngapa   ka   wanti-mi   wajirrkinyi-rla
         water(ABS) PRES   fall-NONPAST   greentime-LOC
         'Rain falls in the "green" season'

These uses of the locative correspond to adjuncts in English.

   A striking difference between the obliques in Warlpiri and English is in
the way in which the argument/adjunct distinction is drawn. Aside from the
occasional idiomatic uses, as with *kardi-ngka* ('card' LOC) 'with cards' in (42),
a Warlpiri semantic case always seems to be usable wherever its meaning would
make sense. ldiosyncratic restrictions such as those discussed for English in
2.3.1 are quite rare. Most usages of Warlpiri oblique cases thus behave like
adjuncts in English. The idiomatic uses might be taken to be P-objects, so there
would be a few representatives of this category, but there seems to be nothing
whose grammatical behaviour corresponds to that of P-complements. It may be
that this impression is a consequence of our insufficient knowledge of Warlpiri,
and that more study might reveal the familiar categories, but at the moment it
seems that in Warlpiri the argument–adjunct distinction is much more closely
aligned with the core–oblique distinction than it is in English.

3   Core grammatical functions

In this section we examine core grammatical functions in detail. As discussed
at the beginning of section 1, core grammatical functions are those express-
ing A, S and P, along with any others that behave like these rather than like

obliques. Core functions are interesting for several reasons. First, they are used to express a wide range of semantic roles beyond the clearcut cases of agent and patient that provide the basis for defining A and P. Furthermore they tend to be syntactically 'active', participating in a wider range of grammatical processes than obliques. Finally, and most interestingly, they are usually (but perhaps not always) associated with what we have called 'grammatical relations': structural relationships, which could plausibly be regarded as structural primitives, which play an important role for the functioning of grammatical principles, but are often abstract with respect to coding features or semantic and pragmatic properties, or both.

The most commonly found and best evidenced grammatical relation is one expressing A and S functions, commonly known as 'subject' (although we shall see that this single label covers at least two rather different kinds of function). But the very prevalence of the subject grammatical relation perhaps leads people to be insufficiently critical in evaluating the evidence for its presence in particular languages. Therefore in 3.1 we will spend considerable time on the issue of how to argue that a subject grammatical relation is present in a language. Then in 3.2 we will look at some of the other grammatical relations that can be argued for in languages that have subjects. An important feature of our approach to subjects is that the evidence does not always support their existence in a language; in the remaining subsections we consider various kinds of languages in which subjects as we have defined them don't exist (although they will show subject-like grammatical relations that we will introduce later). Our conclusion will be that, although a subject grammatical relation does play an important role in the typology of grammatical relations, the subject as traditionally recognized in languages such as English, Latin, and Greek combines two distinct kinds of 'prominence' which in many other languages are kept distinct.

3.1   Subjects

'Subject' is perhaps the oldest grammatical relation concept, found for example in the work of Aristotle.[12] There is furthermore a considerable amount of evidence in different languages for some kind of abstract grammatical relation associated with NPs traditionally regarded as subjects, much more so than for other grammatical relations. But there has unfortunately been considerable flexibility in the use of the term, with concomitant weakening of content, and controversy as to whether subjects are present in various languages. We will here adopt a rather narrow conception of subject, so that it will be relatively easy to assess whether or not we have evidence for the existence of a subject in

---

[12] See Kneale and Kneale (1962) for discussion of the history of the term and concept.

this sense in a particular language (languages without a subject in our present sense might, however, have subjects under some other definition; it's how the content of the definitions applies to particular languages that is interesting, not the terms that are used as labels).

After introducing our concept, we will discuss the various ways in which it can be applied to assess whether or not a subject is present in a language.

3.1.1   A concept of subject   The concept of subject proposed here is that it is a grammatical relation that is the normal expression of the A and S grammatical functions, but not others such as P or obliques. As a grammatical relation, the subject concept should function as a significant ingredient in many of the grammatical phenomena of the language, so that it seems plausible to recognize it as a structural primitive. There are two major kinds of phenomena that are relevant to establishing the existence of subjects: first, the overt coding features in ordinary main clauses; and second, a variety of more complex and abstract grammatical phenomena, such as 'subject ellipsis', coding features in subordinate clauses, and others. When the coding features unambiguously indicate that a subject grammatical relation is present, the more abstract criteria seem invariably to concur. But it is also possible for the coding features to give no evidence, or equivocal evidence, about the presence of a subject. Then the more abstract properties sometimes show that there is a subject, sometimes not. We examine each of these phenomena in turn.

3.1.2   Subjects and coding features in ordinary main clauses   In English and many other languages, there is a variety of coding features in ordinary main clauses that distinguish A of transitive clauses and S of intransitives from P of transitives and other grammatical functions such as obliques. For English, these include nominative as opposed to accusative case, preverbal position, and verb agreement:

(47)   a. He   praises them
              NOM.SG SG   ACC.PL

        b. He   sleeps
              NOM.SG SG

The fact that these and more grammatical phenomena treat A and S alike indicates that, in English, these should not be thought of as independent grammatical functions, but as related ones, most straightfowardly by an analysis in which they are both expressed by a single grammatical relation, which, given our definition, we can call 'subject'.

A great many languages, including many of the familiar modern and classical languages of Europe, follow this pattern of unambiguous evidence for a subject grammatical relation on the basis of some combination of the coding features

of word order, case marking and agreement. In Ancient and Modern Greek, for example, subjects of ordinary main clauses occupy no definite position, but are for the most part regularly marked by nominative case and agreement with the finite verb.

But coding features frequently fail to give a clear indication of grammatical relations, or else give inconsistent indications, as happens, for example, in Warlpiri. We have already seen in 1.2.2 that Warlpiri NP-marking assigns ergative case to NPs with A function, and absolutive to NPs with P or S function. Case marking therefore does not reflect a subject grammatical relation. But the cross-referencing system does.

The NPs that are cross-referenced are those with the cases labelled as 'syntactic' in 1.2.2: ergative, absolutive and dative. Cross-referencing of absolutive and ergative NPs has already been illustrated in example (15) in 1.2.3, repeated below for convenience. Example (48) illustrates cross-referencing of a dative. Example (49) illustrates the failure of cross-referencing to apply with a semantic case, the allative:

(15)
Nya-nyi   ka-rna-palangu   wawirri-jarra   (ngajulu-rlu)
see-NONPAST PRES-1SG(SUBJ)-3DU(OBJ) kangaroo-DU(ABS) (1SG-ERG)
'I see two kangaroos'

(48)   Ngaju ka-rna-ngku   nyuntu-ku wangka-mi
        I(ABS) PRES-1SG(SUBJ)-2SG(OBJ) you-DAT   talk-NONPAST
        'I am talking to you'

(49)   Ngaju ka-rna   nyuntu-kurra parnka-mi
        I(ABS) PRES-1SG(SUBJ) you-ALL   run-NONPAST
        'I am running toward you'

The form of the markers is not determined directly by the case of the NP being cross-referenced. Rather, it seems to be determined primarily by a subject–object distinction in grammatical relations quite similar to that found in English.

There are two sets of cross-reference markers, one for subjects, and another for objects. The cross-referencing is for number (singular, dual, plural) and person (first, second and third), with an inclusive–exclusive distinction in the first person dual and plural (see Bickel and Nichols in vol. III, chapter 3, section 8, for discussion of these inflectional categories), with a limited case-distinction in the object markers. The subject set is used to cross-reference NPs with A or S function, regardless of whether their case is ergative or absolutive:

(50)   a. Ngaju ka-rna   purla-mi
           I-ABS   PRES-1SG(SUBJ) shout-NONPAST
           'I am shouting'

b. Nyuntu   ka-npa   purla-mi
   you(SG.ABS) PRES-2SG(SUBJ) shout-NONPAST
   'you are shouting'

c. Ngajulu-rlu ka-rna   yankirri   wajilipi-nyi
   I-ERG   PRES-1SG(SUBJ) emu(ABS)   chase-NONPAST
   'I am chasing an emu'

The object markers cross-reference NPs with P function, which are absolutive, and also NPs in the dative case. Examples of absolutive object cross-referencing are:

(51) a. Ngarrka-ngku ka-ju   ngaju   panti-rni
         man-ERG   PRES-1SG(OBJ) I(ABS)   spear-NONPAST
         'The man is spearing me'

      b. Ngaju   ka-npa-ju   nyuntulu-rlu nya-nyi
         me(ABS)   PRES-2SG(SUBJ)-1SG(OBJ)   you-ERG   see-NONPAST
         'You see me'

      c. Ngajulu-rlu ka-rna-ngku   nyuntu   nya-nyi
         I-ERG   PRES-1SG(SUBJ)-2SG(OBJ)   you(SG.ABS)   see-NONPAST
         'I see you'

Dative objects are cross-referenced by the same markers as are used for absolutives, except in the third person singular where *-rla* is used instead of zero. Dative objects will be discussed in 3.2.2

There are various additional principles which determine the form of cross-referencing in examples more complex than these (such as those involving plurals). These are described in great detail in Hale (1973), and needn't be considered here. But these complications do not alter the basic point that the systems of NP-marking and cross-referencing give conflicting testimony as to what the basic grammatical relations of A, S and P are.

Therefore, coding features do not always provide consistent indications for grammatical relations. Does this mean that the grammatical relations are present, but coded inconsistently, or simply absent from the languages in question? The situation for each language should be decided on its merits. For some languages, reasonable cases have been made that grammatical relations such as subject and object are absent (Bhat (1991)), but in others, such as Warlpiri, other aspects of grammatical behaviour seem to indicate that they are present.

3.1.3   Subject ellipsis   Perhaps the commonest property of subjects that is useful for identifying them is their tendency to be optionally or obligatorily ellipsed in various kinds of grammmatical constructions, especially

multi-clause sentence structures. A highly typical example from English is provided by adverbial clauses introduced by the conjunction *while*.

These clauses take two forms. In one, *while* is followed by an ordinary clause structure with a subject and a tensed verb. In the other, the subject is omitted and the verb put in the (gerund) *-ing* form, which does not show agreement:

(52)   a. The student watched the guard while he killed the prisoner
         b. The student watched the guard while killing the prisoner

When the verb is tensed, the subject must be included; when the verb is in the *-ing* (gerund) form, its subject must be omitted, but is understood as being the same as the subject of the main clause:

(53)   a. *The student watched the guard while killed the prisoner.
         b. *The student watched the guard while he/his/him killing the
             prisoner

Omission of a non-subject NP will not satisfy the requirement, as the reader can easily verify. The subject relation thus functions in the principles governing the form of *while*-constructions.

It is also involved in a principle governing their interpretation. In (52a) we could understand the *while*-clause subject as referring to the guard, the student, or some third person. In the absence of wider context, we tend to interpret it as referring to some NP within the sentence, and from our knowledge of the world we tend to assume that it refers to the guard rather than to the student.

But the interpretation of (52b) is not so free. Here we would normally understand the student rather than the guard to be killing the prisoner, in spite of the oddity of this situation. There seems to be a principle to the effect that a *while* + gerund construction is interpreted as if it had a subject coreferential to the subject of the matrix clause (note that if *while* is omitted, we immediately understand the subject of the gerund to be coreferential with the object rather than the subject of the matrix clause).

On the basis of (52, 53) alone, one might venture an alternative account, in which it is the agent rather than the subject of the *while* + gerund construction that is suppressed, and the subject of the *while* construction is understood as being the same as the agent rather than the subject of the main clause. In this kind of account, we would have a direct connection between the overt form and the meaning, without an intervening level of grammatical relations.

This possibility may be discounted on the basis of sentences such as *John felt apprehensive while being wheeled into the operating room*, in which the overt and 'understood' subjects are not agents, and even more strongly by examples in which the *while* + gerund construction is combined with the passive construction:

(54)   a. The student watched the guard while killing the prisoner
        b. The student watched the guard while being killed by the prisoner
        c. The student was watched by the guard while killing the prisoner
        d. The student was watched by the guard while being killed by the
           prisoner

It is the subject of the matrix that is understood as the subject of the gerund, regardless of the semantic roles involved, and of how unusual the situation described is.

It also seems that no well-defined pragmatic notion such as topicality is the conditioning factor, although this is hard to show conclusively, since pragmatic functions are generally more elusive and less well understood. For example, in a sentence such as *A guard tortured the prisoner while watching television*, it seems pretty clear that *the prisoner* can be the topic. Nonetheless, the principle for the interpretation of the *while* + gerund construction continues to operate as before.

Phenomena such as these illustrate the need for a level of syntactic structure at which abstract grammatical relations such as subject are defined, which are distinct from semiotic concepts, and which are significant for the functioning of grammatical rules.

From a theoretical point of view, there are three major possibilities for the analysis of *while* + gerund constructions. The first is that the gerund has no subject in syntactic structure, but that the principles of semantic interpretation treat it as if it had a subject coreferential with that of the matrix (main) clause. Second, the gerund might have a subject in the syntactic structure which is coreferential with the matrix subject, but which does not appear in the overt form of the sentence. The third possibility is that the theory of sentence structure characterizes the NP in matrix subject position as the subject of both the main clause and the gerund.

The choice between these possibilities is a complicated question, which does not concern us here. What matters here is that, whatever approach is taken, it is clear that the notion of subject plays a central and obvious role in the description of the constructions: it is the subject of the subordinate clause that is obligatorily omitted, and the subject of the matrix that obligatorily serves as its 'controller', that is, as the NP that is understood as the subject of the subordinate clause.

Subject ellipsis can often be used to provide more evidence about grammatical relations when the coding features are equivocal. In Warlpiri, there are counterparts to the *while* + gerund construction that show that this language has a subject grammatical relation (one expressing s and a functions) in spite of the inconsistent testimony of the coding features. These are 'infinitival' subordinate clauses (adverbial or relative in sense), in which no auxiliary appears, but an 'infinitival complementizer' is attached to the verb, which then appears finally

in the infinitival phrase, and can't be reordered within it (there is, however, a possibility of nominals within the infinitive phrase 'leaking' out of it into the matrix (Laughren (1989))).

Many of the infinitival complementizers require suppression of the complement subject, imposing various conditions on what it may be understood to be coreferential with. One of these is the complementizer *kurra*, which expresses action simultaneous with that of the main verb, and imposes the condition that the complement subject be coreferential with a non-subject (preferably object) argument of the matrix:

(55)

a. Ngajulu-rlu-rna yankirri pantu-rnu, ngapa nga-rninja-kurra
   I-ERG-1SG(SUBJ) emu(ABS) spear-PAST water(ABS) drink-INF-while
   'I speared the emu while it (not I) was drinking water'

b. Ngarrka-rna nya-ngu wawirri panti-rninja-kurra
   man(ABS)-1SG(SUBJ) see-PAST kangaroo(ABS) spear-INF-while
   'I saw a man spear a kangaroo'

c. Ngaju ka-rna-ngku marri-jarri-mi nyuntu-ku
   I(ABS) PRES-1SG(SUBJ)-2SG(OBJ) grief-being-NONPAST you-DAT
   murrumurru nguna-nja-kurra(-ku)
   sick lie-INF-while(-DAT)
   'I feel sorry for you while you are lying sick'

d. Karli-rna nya-ngu pirli-ngirli wanti-nja-kurra
   boomerang(ABS)-1SG(SUBJ) see-PAST stone-ELATIVE fall-INF-while
   'I saw the boomerang fall from the stone'

The infinitival verbs of (55a–b) would take ergative subjects if finite, those of (55c–d) absolute. The examples also illustrate a variety of semantic roles for the omitted subject and its controller.

It is crucial to the argument that *kurra* requires (rather than merely permits) omission of the subject: since NPs can be rather freely omitted in Warlpiri, if a complementizer merely permits an omitted argument in its clause to be understood as coreferential with one in the matrix, without actually *requiring* omission and understood coreference, we could simply say that the omitted argument was an ellipsed anaphoric pronoun which happened to be coreferential with an NP in the matrix clause (this would often be permitted by the usual principles governing null anaphora). There would then be no syntactic phenomenon specifically associated with the subject of a -*kurra* complement. The more general point is that what needs to be shown is some *difference* in omissibility from ordinary clauses. In English, for example, NPs aren't freely omissible, so the possibility of omission in the *while* + gerund construction is enough to make an argument for a grammatical relation, whereas in Warlpiri

or other languages where NP omission is widespread, something stronger is required, such as obligatory omission, and it must furthermore not be possible to describe the class of NPs which can be omitted in purely semantic terms.

It is also important that the phenomenon involves a variety of semantic roles. If, for example, only agents were obligatorily suppressed in this construction, one could claim that the principle referred to agent, a semantic role, rather than to a grammatical relation in sentence structure.

The Warlpiri and English constructions we have discussed so far are both adverbial in nature, but subject ellipsis can be an optional or obligatory feature of virtually any kind of subordinate or coordinate clause construction in some languages. English nonfinite (participial) relative clauses require ellipsis of the subject (which is understood as coreferential with the head):

(56)   a. People [reporting their neighbours to the authorities] will be
           rewarded.
        b. People [reported by their neighbours to the authorities] will be
           investigated.

(57)   a. *People [their neighbours(') reporting] will be investigated
        b. *People [their neighbours(') reported by] will be investigated

English has other strategies of relativization that can be used on non-subjects, but, in some languages, such as Malagasy (Keenan and Comrie (1977)), relativization is possible only for subjects.

Complement and coordinate clauses can also be useful in arguing for a subject grammatical relation, although probably not as often as adverbial clauses. We present an example from Icelandic. In this language, there is a subject clearly identified by the coding features of preverbal position (in sentences without topicalization), nominative case, and agreement with the verb:

(58)   a. Við   dönsuð-um
           we(NOM.IPL) danced-1PL
           'We danced'

        b. Þeir   dó-u
           they(NOM.PL) died-3PL
           'They died'

        c. Þeir   drápu   hunda-na
           they(NOM.PL) killed-3PL dogs-the(ACC)
           'They killed the dogs'

But there is also a group of verbs that take an NP in the regular preverbal 'subject' position, but show differences in other coding features – the case can be genitive, dative or accusative, and there is no verb agreement:[13]

(59)   a. Đá   vantar   peninga
          them(ACC.PL) lacks(3SG) money(ACC)
          'They lack money'

       b. Mér   líkar   vel   við   henni
          me(DAT)   likes(3SG) well with her(DAT)
          'I like her'

Complement subject ellipsis provides evidence that these non-nominative NPs are subjects in spite of lacking the coding features of nominative case and agreement.

This is provided by the considerable number of verbs taking infinitival complements introduced by the complementizer *að*. These complements require their subject to be missing, and understood as coreferent to the main clause subject:

(60)   a. Ég   vonast til   að sjá hana
          I(NOM)   hope   toward   to see her(ACC)

       b. *Ég   vonast til   að ég sjá hana
          I(NOM)   hope   toward   to I   see her(ACC)

       c. *Ég   vonast til   ég að sjá hana
          I(NOM)   hope   toward   I   to see her(ACC)
          'I hope to see her'

Examples (b) and (c) are bad because they contain attempted complement subjects in position before and after *að*, which is not possible.

But the putative oblique subjects like those of (59) do basically satisfy the requirement that a subject be ellipsed, although ellipsis of a non-nominative subject does result in some degradation of acceptability (Thráinsson (1979:301–4, 469); Andrews (1990)):

(61)   Ég   vonast til   að vanta ekki peninga
       I(NOM)   hope   towards   to lack not money
       'I hope not to lack money'

This example also illustrates that the process applies to more semantic roles than just agents.

---

[13] There is considerably more to the subject position than just a tendency to appear first. See Jónsson (1996) for a recent analysis of Icelandic clause structure.

Coordinate structures provide another possibility. In many languages, when clauses are conjoined, it is possible to omit an NP in one conjunct if it is coreferential with one in another conjunct, and if the NPs have the same grammatical relation in their respective conjuncts. Icelandic is one of those languages. In (62a), we find that an oblique subject of a coordinated clause may be omitted under coreference with the subject of a preceding conjunct (Rögnvaldsson (1982)), while in (62b) we see that this is not possible for an object:

(62)   a. Ég   sá   stúlkuna   og Ø líkaði vel við henni
           I(NOM) saw the girl(ACC) and [I] liked well with her(DAT)
           'I saw the girl and liked her'

       b. *Ég   sá   stúlkuna   og hún   heyrði Ø
           I(NOM) saw the.girl(ACC) and she(NOM) heard [me]
           'I saw the girl and she heard me'

(Objects can however be omitted upon coreference to preceding objects, see Thráinsson (1979:471).) Likewise, only a subject, including oblique subjects, may control the ellipsis of the subject of a coordinate clause:

(63)   a. Þeim   líkar maturinn   og Ø   borða mikið
           them(DAT) likes the food(NOM) and [they] eat   a.lot
           'They like the food and eat a lot'

       b. *Þeir   sjá stúlkuna   og Ø   heyrir þá
           they(NOM) see the.girl(ACC) and [she] hears them
           'They (masc.) see the girl and she hears them'

As with the other kinds of instances of subject ellipsis, it is necessary to ascertain that there isn't any free and general process of NP ellipsis that might be responsible for the 'missing subjects' in order for there to be evidence of a subject grammatical relation, and that its conditions can't be described in purely semantic terms.

3.1.4   Coding features in non-main clauses   It frequently happens that the coding features of subjects are different in subordinate clauses from those in main clauses. One of the commonest instances of this is when subjects of subordinate clauses acquire special case marking. In English, for example, the subject of a gerund can be accusative or genitive, but not nominative, which is the normal case for subjects:

(64)   a. Him/*he running Ewing Oil is difficult to imagine
       b. His/*he running Ewing Oil would upset a lot of people

Another is the Ancient Greek 'circumstantial participle' construction discussed in 1.2.3, with examples (16) and (17). If the subject of the participle is not

coreferential with any NP in the matrix, it is expressed in the genitive instead of the nominative, which is the normal case for subjects:

(65) Ape:nte:sa Philippo:i   Klearchou   apiontos
      I.met   Philip(DAT)   Klearchus(GEN) leaving(GEN)
      'I met Philip while Klearchus was leaving'

Cross-referencing is also affected: finite verbs in Greek agree with their subjects in person and number, while participles agree in gender, number and case (but infinitives, which take accusative subjects, don't agree at all).

   Special NP-marking in subordinate clauses is usually restricted to subjects, although it sometimes involves other core grammatical relations such as object, as for example in the Saibai dialect of Kala Lagaw Ya (Comrie (1981)). Sometimes subordinate-clause coding features provide useful arguments for subjecthood. This happens in Warlpiri. For older speakers, instead of requiring subject ellipsis, some nonfinite clause constructions permit the subject to be expressed, and some of these permit or require a special case-marker on that subject (Nash (1980:233–4)). One of these complementizers is *-rlarni*, whose meaning specifies that the action of the complement is contemporaneous with that of the matrix. Below are some examples with this complementizer:

(66) Ngarrka-ngku-ka karli   jarnti-rni . . .
      man-ERG-PRES   boomerang(ABS) carve-NONPAST
      'The man is carving the boomerang, . . .'

   a. . . . kurdu-ku/-Ø   purla-nyja-rlarni
            child-DAT/(ABS)   shout-INF-while
         'while the child is shouting'

   b. . . . kurdu-ku/-ngku maliki   wajilipi-nyja-rlarni
            child-DAT/ERG   dog(ABS)   chase-INF-while
         'while the child is chasing the dog'

   c. . . . karnta-ku/-ngku   kurdu-ku   miyi   yi-nyja-rlarni
            woman-DAT/-ERG   child-DAT food(ABS) give-INF-while
         'while the woman is giving food to the child'

The subject takes either its normal case marking or the dative. Furthermore the subject, if it is there, must be initial in the *-rlarni* complement, regardless of its case marking. If, for example, *kurdu-ku* were placed after *maliki* in (66b) above, it would have to be interpreted as a Beneficiary, so the meaning would be 'The man is carving the boomerang, while somebody is chasing the dog for the child' (Laughren (p.c.)).

   The *-rlarni* construction, in sharp contrast to main clause constructions, expresses the subject grammatical relation directly in terms of both case-marking and linear ordering: the subject may be marked dative instead of its

usual case (regardless of whether that is ergative or absolutive), and the subject must be initial in the complement. These additional phenomena complete the case for the existence of subjects in Warlpiri.

3.1.5   Switch reference   The third grammatical test for identifying subjects that we will discuss involves what are called 'switch reference' systems. These are systems in which the verb of a clause bears a marker which indicates, among other things, whether the subject of that clause is the same or different from that of some other coordinated or subordinated clause.

P. K. Austin (1981a, 1981b) uses switch reference to argue for subjects in the Australian language Diyari. Grammatical relations are not directly reflected by coding features in Diyari because, like many other Australian languages, Diyari has a 'split ergative' case-marking system in which different sorts of nominals have different systems of case forms for A, S and P. First and second person non-singular (dual and plural) pronouns have a nominative (A/S) and accusative (P); singular common nouns and masculine proper names have an ergative (A) and an absolutive (P/S), while all other nominals have distinct forms for all three functions: ergative (A), absolutive (S) and accusative (P).

Most complex sentence constructions have switch reference marking expressed as an affix on the verb of the subordinate clause. The affix indicates the type of construction, and whether the subjects of the two clauses are the same or different. One of these constructions is the 'relative clause', a type of subordinate clause which further specifies either some participant in the main clause (an 'NP-relative' interpretation (Hale (1976))), or the time of the clause (a 'T-relative' interpretation). If the subject (A or S NP) of the subordinate clause is the same as that of the main clause, *-na* is added to its verb: if the subjects are different *-ṇaṇi* is added. It is the A/S function rather than the case forms that is relevant for the switch reference system.

Example (67) is an assortment of subordinate clauses with same subject (SS) marking, (68) an assortment with different subject (DS) marking. Note that the subordinate clause corresponds to a considerable range of subordinate clause types in English, including relative clauses, *when*-clauses, conditionals, and complement clauses. A shared subject may or may not be deleted in the subordinate clause.

(67)

a. ṇawu   ṭika-ṇa   / ṇawu   yaṭa-l̠   ṇana-yi   yiŋaṇu
    he(ABS)   return-REL(SS) he(ABS) speak-FUT   AUX-PRES   you(SG.LOC)
    'If he comes back he'll talk to you'

b. ŋaṯu   kanṯa   kulʸakulʸa ṯayi-ṇa   / ŋaṇi   piṯi-yi
    I(ERG)   grass(ABS) green(ABS) eat-REL(SS)   I(ABS)   fart-PRES
    'When I eat green grass, I fart'

c. winṯa ŋani   pali-ṇa   / ŋaṯu kaṇa   ŋakaṇi   ŋama-1ka-yi ṇaka
   when I(ABS) die-REL(SS) I(ERG) person me(DAT) sit-TR-PRES there(LOC)
   'When I die, I will have my people there'
   (ŋakaṇi is here functioning as a possessive modifier of kaṇa 'person')

(68)
a. kan^yt^yi mindi-ya ŋani   / ṇaka-lda ṇawu   wakaṛa-ṇani
   can   run-PAST she(ABS) there-LOC he(ABS) come-REL(DS)
   'She could have run (the distance) if he had come back again'

b. ṯanali   nina   nayi-yi / nina   waraṛa-ṇa wanti-ṇani
   they(PL.ERG)   he(ACC) see-PRES he(ACC) leave-PART AUX-REL(DS)
   'They see him after he had been left (for a long time)'

c. ŋani   niŋki-ya   wakaṛa-ṇa   / ŋaṯu nana   wiḻa
   I(ABS)   here(LOC)   come-REL(SS) I(ERG) she(ACC) woman(ABS)

   nayi-yi / yinda-ṇani
   see-PRES cry-REL(DS)
   'When I come here I see that woman [who is] crying'

In (67a), there are coreferential NPs in s function in the two clauses, so ss-marking appears. In (67b) the main clause s is coreferential with the (preceding) relative clause A, so again ss-marking appears, even though the coreferential nominals differ in their case forms. In (67c), the relative clause s is coreferential with the matrix A, so again ss marking appears.

In (68a), the matrix and relative clause contain no coreferential NPs, so DS-marking appears. In (68b) there are coreferential NPs, but they are Ps in both clauses (it is understood that the people who see him are different from the ones who left him, who are represented by an ellipsed subject for the clause). In (68c) there are two relative clauses, the first with a temporal interpretation with s coreferential with the matrix A, the second interpreted as a perception complement with s coreferential with matrix P. So the first relative clause has ss-marking, the second DS-marking.

Switch reference in this and other types of subordinate clauses provides evidence that Diyari has a subject grammatical relation comprising A and s functions, in spite of the completely ambiguous testimony of the NP-marking system. More than simple coreference between subjects is normally involved in switch reference systems; see Stirling (1993) for a detailed study.

3.1.6   Reflexivization Many languages have special pronouns, called reflexive pronouns, that are used to indicate that an NP is coreferential with an NP bearing a certain structural relationship to it. In many languages, such pronouns are used when an NP is to be coreferential with the subject of a clause that contains it.

One such language is Malayalam (Mohanan (1982)). Malayalam has free word order, expressing grammatical relations by NP marking. NPs in A/S function are nominative, Ps are accusative if animate, nominative if inanimate. In this language, the reflexive possessive pronoun *swaṇṭam* requires an antecedent which is a subject (either of the clause immediately containing *swaṇṭam*, or of some higher one). Therefore the following two sentences are good, even though *swaṇṭam* follows its antecedent in the first and precedes it in the second, since in both cases the antecedent is subject:

(69)   a. Raajaawə   swaṇṭam bhaaryaye ṇulli
           king(NOM) self's   wife(ACC)   pinched

        b. Swaṇṭam bhaaryaye raajaawə   ṇulli
           self's   wife(ACC)   king(NOM) pinched
           'The king pinched his own wife'

But when an attempted antecedent is an object, the result is ungrammatical:

(70)   *Raajaawine swaṇṭam bhaarya   ṇulli
          king(ACC)   self's   wife(NOM) pinched
          'His own wife pinched the king'

These examples show that reflexivization depends on the grammatical relations rather than on linear order.

Like English, Malayalam has a passive construction in which the argument expressed as an object in the active is expressed as the subject, and the argument expressed as the subject in the active is expressed as an instrumental (with the ending *-aal*). The interaction of reflexivization with passivization shows it to be dependent on grammatical relations rather than semantic roles such as agent and patient. The controller of the reflexive has to be the subject even when in the passive construction the subject is the patient:

(71)   a. Raajaawə   swaṇṭam bhaaryaal   ṇullappeṭṭu
           king(NOM) self's   wife(INSTR)   pinch(PAST.PASS)

        b. Swaṇṭam bhaaryaal   raajaawə   ṇullappeṭṭu
           self's   wife(INSTR)   king(NOM) pinch(PAST.PASS)
           'The king was pinched by his own wife'

(72)   *Raajaawinaal swaṇṭam bhaarya   ṇullappeṭṭu
          king(INSTR)   self's   wife(NOM) pinch(PAST.PASS)
          'His own wife was pinched by the king'

Malayalam shares with Icelandic the feature of having NPs that lack some of the usual coding properties of subjects, but show some of their other grammatical properties, as do many other of the languages of South Asia (Masica (1971); Verma and Mohanan (1990)). In Malayalam, most verbs take subjects in the

nominative (unmarked) case, but some seem to have subjects in the dative case as a lexical property. Also certain derivational affixes, such as the desiderative -*aṇam*, impose the requirement that the derived verb take a dative subject (if a verb v means 'to X', the verb V-*aṇam* means 'to want to X'). Reflexivization provides one of the arguments that these datives are indeed subjects, since they can antecede *swaṇṭam*, while dative recipients with ordinary verbs of giving cannot:

(73) a. Raajaawinə swaṇṭam bhaaryaye n̪ull-aṇam
          king(DAT)   self's   wife(DAT) pinch-DESIDERATIVE
          'The king wants to pinch his wife'

      b. Raajaawinə swaṇṭam bhaaryaye ṣṭam-aanə
          king(DAT)   self's   wife(ACC) liking-is
          'The king likes his wife'

      c. *Raajaawə makalikkə   swaṇṭam bhartaawine   koṭuttu
          king(NOM) daughter(DAT) self's   husband(ACC) give(PAST)
          'The king gave his daughter her husband'

In a similar fashion, reflexivization also provides evidence for dative subjects in various other South Asian languages, such as Hindi (Kachru, Kachru, and Bhatia (1976)).

3.1.7   Other properties of subjects   There are a very large number of other properties that subjects can have in a language, too many to list here. In Icelandic, for example, there are currently at least thirteen known properties that can be used to argue that certain non-nominative NPs are subjects (Andrews (2001)). An early compilation of common subject properties is Keenan (1976c); see Manning (1996:12–14, 17) for more recent discussion. The most important point is that it is not sufficient simply to note that some property that frequently characterizes subjects in other languages happens to be true of subjects in the language under discussion: it must also be shown that the property does not apply to non-subjects, and that it cannot be described solely in terms of semantic roles.

For example, in English, one can note that reflexive pronouns can have the preverbal NP as their antecedent, in the same way that Malayalam reflexive pronouns can have the nominative as their antecedent:

(74) a. John$_i$ talked about himself$_i$
      b. John$_i$ told Mary$_j$ about himself$_i$

But there is no argument for subjects in English here, because non-subjects can also be the antecedent of reflexive pronouns:

(75)   John$_i$ told Mary$_j$ about herself$_j$

To provide evidence for a grammatical relation of subjects, a property must apply to the putative subjects but not to clear cases of non-subjects, and most also not be statable in terms of other concepts such as semantic roles, since we shouldn't postulate abstract concepts such as grammatical relations if other independently motivated concepts are sufficient to account for the phenomena.

3.2   *Other core grammatical relations*

In this subsection we discuss some of the other core grammatical relations that are commonly found in languages that have subjects. These grammatical relations are commonly called 'objects': direct objects, indirect objects, and so forth. Objects are generally more problematic than subjects because there are fewer grammatical processes applying exclusively to specific types of objects. It can therefore be difficult to tell whether variations in the coding features of object-like NPs reflect differences in their grammatical relations. Some important studies and collections on aspects of objecthood are Plank (1984), Dryer (1986), M. C. Baker (1988), and Alsina (1996a).

The most important type of object, and the most widely distributed, is the direct object. These are discussed immediately below, together with the highly similar second objects. Next we consider indirect objects, and then finally certain other less commonly found core grammatical relations.

3.2.1   *Direct objects and second objects*   We have already defined 'direct object' as the grammatical relation, if there is one, associated with P function. There turn out to be two potential kinds of problems that arise in connection with recognizing direct objects. The first is that sometimes P function is expressed by more than one morphosyntactic technique, without there being a clear basis for saying that there is a difference in grammatical relations. Most commonly, animate and/or definite Ps are expressed differently from inanimate and/or indefinite ones. In Hindi, for example, animate Ps require the accusative case-marker *ko*, while inanimate Ps allow (but don't require) the marker if they are definite, and don't allow it if they are indefinite (T. Mohanan (1994:79–80)). Besides accusative, the inanimate objects can be nominative if they are definite, and must be if they are indefinite:

(76)   a.   Ilaa ne   bacce-ko/*baccaa   uṭaayaa
             Ilaa ERG child-ACC/child(NOM) lift(PERF)
             'Ilaa lifted the/a child'

        b.   Ilaa ne   haar   uṭaayaa
             Ilaa ERG necklace(NOM) lift(PERF)
             'Ilaa lifted the/a necklace'

   c. Ilaa ne   haar-ko   uṭaayaa
      Ilaa ERG necklace-ACC lift(PERF)
      'Ilaa lifted the/*a necklace'

A somewhat similar phenomenon appears in Spanish. Here full NP animate objects, regardless of definiteness, are marked with an object-marker *a* (77a), while inanimates are marked by nothing (77b). But pronominal objects are marked by an accusative case clitic in front of the verb, regardless of animacy:[14]

(77)   a. Vimos   a   alguien
         we.saw OM somebody
         'We saw somebody'

       b. Vimos   (*a) el   interruptor
          we.saw (OM) the   switch
          'We saw the switch'

       c. Lo   vimos
          it(ACC) we.saw
          'We saw it/him'

Pronominal animate objects can also be 'doubled' as full NPs, in which case one sees the marker on the NP together with an accusative pronoun:

(78)   Lo   vimos   a   él
       him(ACC) we.saw   OM him
       'We saw HIM'

There doesn't seem to be any solid basis for saying that one or the other of these treatments is characteristic of a 'real P' (participant receiving the normal treatment accorded to a patient of a PTV). Rather, in many languages there are just two treatments, apportioned in accord with animacy, or definiteness. This is in fact probably the commonest situation in which there are two different ways of expressing P function. It is important that in all such cases, all PTVs can use either technique, providing that its semantic/pragmatic conditions are met.

   A different but related issue is whether or not there are two grammatical relations involved. In the case of Spanish, this seems unlikely: both animate (marked) and inanimate (unmarked) P are represented the same way, as accusative clitics, when pronominal, and both can likewise be passivized. In the case of Hindi, there doesn't seem to be a comparable argument for identifying the two treatments of P as one grammatical relation, but neither is there any against it, and treating them as the same gives us a simpler account of verbs in

---

[14] There is however an option, called *leismo*, of using dative rather than accusative forms of animate object pronouns, so the treatment of these two types seems to be diverging.

the lexicon, since a transitive verb will simply be specified as taking a subject and an object, rather than a subject and one of two types of object. So one can say that P is consistently realized by a grammatical relation, although the evidence for this is not overwhelming.

The second problem is more serious, which is that of distinguishing Ps from potential cases of non-Ps. These cases arise in at least two ways. First, there can be 'non-canonical' objects that share some but not all of the properties of P. Second, there can be 'multiple' objects where there is more than one NP that shows some of the characteristic properties of P.

The first kind of case often arises in languages where grammatical relations are coded by NP-marking. In such languages, it often happens that a large number of two-argument verbs take non-subject arguments in some case not normally found on P. In Warlpiri, we have noted verbs taking non-subject arguments in the dative and locative cases (examples (48) and (42), respectively). Simpson (1991:311–17) argues that these dative arguments should be considered as objects because they can be cross-referenced like ordinary objects, and serve as controllers for *kurra* nonfinite clauses, as illustrated in (55c).

Another kind of example is afforded by German. Here Ps are expressed as accusative NPs, illustrated in (79a). But there are a fair number of two-argument verbs that take their second (non-subject) argument in the dative, illustrated in (b):

(79) a. Sie   sah   ihn
          she(NOM) saw   him(ACC)
          'She saw him'
       b. Sie   hilfte ihm
          she(NOM) helped him(DAT)
          'She helped him'.

In German, there don't seem to be any phenomena which clearly unite the accusative of (a) and the dative of (b) as bearers of a single grammatical relation, other than that of appearing as a bare NP, without a preposition.

For example, both kinds of verbs can passivize, but an accusative object becomes nominative and obligatorily occupies the subject position, while the dative retains its dative case, and remains in the VP:

(80) a. Er   wurde   gesehen
          he(NOM)   became   seen
          'He was seen'
       b. *Es   wurde   ihn/er   gesehen
          It(NOM)   became   him(ACC/NOM)   seen
          'He was seen'

    c. *Er   wurde   geholfen
       He(NOM) became helped
       'He was helped'

    d. Es   wurde   ihm   geholfen
       It(NOM) became him(DAT) helped
       'He was helped'

Es 'it' in (d) is functioning as a 'filler' in sentence-initial position in cases where there is no subject; it is impossible in (b) because the passive verb form *gesehen* in this example has the nominative *er* 'he' as its subject.

   It is possible to put the dative into sentence-initial position (like almost any other constituent of the clause), with consequent disappearance of *es*, but these datives pass none of the relevant tests for subjecthood. For example they can't be ellipsed as understood subjects of complements:

(81)   a. Uns   wurde   von der Polizei geholfen
           we(DAT) became by   the police helped
           'We were helped by the police'

        b. *Wir   möchten von der Polizei geholfen werden
           We(NOM) want   by   the police helped   to.become
           'We want to be helped by the police'   Jónsson (1996:127–9)

   By contrast, in Icelandic, when such postverbal dative putative objects are passivized, they obligatorily occupy subject position and pass tests for subjecthood:

(82)   a. Ðeir hálpuðu honum
           they helped   him(DAT)
           'They helped him'

        b. Honum   var hjálpað
           him(DAT) was helped
           'He was helped'

        c. Hann vonast til   að verða hjálpað
           he   hopes toward to be   helped
           'He hopes to be helped'

So in Icelandic we have a reason for grouping the canonically marked (accusative) objects with the non-canonically marked (dative) putative objects, but in German we don't appear to. It may thus require substantial investigation to work out which NPs are direct objects in languages with rich case-marking systems.

The other case tends to arise in systems that code grammatical relations by order. Here it is frequent for two non-subject arguments to appear without distinguishing role-markers, in what is often called a 'double object' construction such as that of *Susan handed Paul the shovel*, mentioned in (35a) above. In this construction, after the verb *handed* appear two bare NPs, *Paul* and *the shovel*. In traditional terminology, *Paul* would be described as the 'indirect object' and *the shovel* as the 'direct object', but this classification is based on the semantic roles, and is partly based on the fact that in many languages with case marking, the recipient would be in the dative case and the theme in the accusative.

Examining a range of languages with double object constructions reveals a rather complex situation. In the most straightforward type, one of the two NPs, usually but not always the one expressing the recipient, takes on all of the grammatical properties of a P, and may thus be non-controversially considered to be the direct object and bearer of P-function.[15]

A language of this type is the Bantu language Chi Mwi:ni (Kisseberth and Abasheikh (1977)). The general form of Chi Mwi:ni sentence structure is not unlike that of English: subjects and objects being unmarked and appearing in SVO order, followed by obliques with prepositional NP marking. There is furthermore a passive construction like that of Malayalam, which puts an extra affix on the verb but does not add an auxiliary. Among the differences is that Chi Mwi:ni has a rich agreement system, with subjects triggering obligatory and objects optional cross-referencing on the verb.

PS are distinguished from SS and obliques by the two properties of triggering optional cross-referencing on the verb (the cross-reference marker appearing between the tense marker, if there is an overt one, and the stem, unlike the obligatory subject cross-reference marker, which precedes the tense marker) and being able to undergo passivization. These two properties are illustrated below:

(83)   a. Nu:ru Ø-Ø-chi-łes-ełe   chibu:ku
           Nuru   he(SUBJ)-PAST-it(OBJ)-bring-ASP book
           'Nuru brought the book'

        b. Chibu:ku chi-Ø-łes-el-a   na Nu:ru
           book   it(SUBJ)-PAST-bring-ASP-PASS by Nuru
           'The book was brought by Nuru'

                              Kisseberth and Abasheikh (1977:192–3)

There are double object constructions in which two NPs appear postverbally without NP marking. The simplest constructions of this sort occur with verbs taking a theme and a goal/source (which may be a recipient or loser, or simply

---

[15] This is sometimes called the 'Primary Object', in part to avert potential confusion with the traditional usage of the term *direct object*, but for theoretical reasons as well (Dryer (1986)).

something to which something is applied, such as a cart that is oiled). We will refer to these non-theme arguments as 'recipients', though their range of semantic roles is wider than indicated by this term.

In a double object construction, both of the properties characteristic of P accrue to the recipient (which normally occupies the immediately postverbal position), as illustrated in (84), but not to the theme, as illustrated in (85):

(84)   a. Nu:ru Ø-m-ɬet-el-ele   mwa:limu chibu:ku
          Nuru he(SUBJ)-him(OBJ)-bring-DAT-ASP teacher   book
          'Nuru brought the book to the teacher'

       b. Mwa:limu Ø-ɬet-el-el-a   chibu:ku na Nu:ru
          teacher   he(SUBJ)-bring-DAT-ASP-PASS book   by Nu:ru
          'The teacher was brought the book by Nuru'

(85)   a. *Nu:ru Ø-chi-ɬet-el-ele   mwa:limu chibu:ku
          Nuru   he(SUBJ)-it(OBJ)-bring-DAT-ASP teacher book
          'Nuru brought the book to the teacher'

       b. *Chibu:ku chi-ɬet-el-el-a   mwa:limu na Nu:ru
          book   he(SUBJ)-bring-DAT-ASP-PASS teacher by Nuru
          'The book was brought (to) the teacher by Nuru'
                             Kisseberth and Abasheikh (1977:192–3)

Note that the presence of a recipient object is signalled by the affix *el* glossed DAT (it is generally called the 'applied' affix in Bantu linguistics). This is not a cross-reference affix because it doesn't show agreement with the grammatical features of the recipient; rather it signals the application of a valence-change operation.

It seems quite unproblematic to view the recipient in the Chi-Mwi:ni double object construction as the syntactic direct object, since it monopolizes the properties of a sole object. The theme in these constructions would then bear a different grammatical relation, which we may call 'second object', or 'secondary object', if the term 'primary object' is being used. The availability in Universal Grammar of a direct vs second object distinction is further indicated by the existence in some languages such as Ojibwa (Rhodes (1990)) of a distinction between normal transitive verbs, which take a subject and a direct object, and 'pseudo-transitive' verbs, which can be strongly argued to take subject and a second object, the same grammatical relation that expresses the theme in a ditransitive verb, the recipient being expressed as a direct (or 'primary') object.[16]

---

[16] True PTVs belong to the normal transitive class, however certain verbs with theme objects are pseudo-transitive. This is an example of why Patients need to be distinguished from themes in defining the class of PTVs.

Chi-Mwi:ni illustrates what is called 'asymmetric' behaviour with respect to object properties: the clause contains multiple NPs whose appearance is not distinct from an NP in P function. Only one exhibits the properties normally exhibited by a sole P (but not an S, and therefore can be plausibly analysed as bearing a 'direct object' grammatical relation). Asymmetric behaviour is widely found in the languages of the world – see, for example, Chung (1976) for an example involving five object properties in Bahasa Indonesia, but there are two additional possibilities.

One is 'symmetrical' behaviour, where more than one of the NPs that superficially look like P also share substantial grammatical behaviour with a sole P. Ojibwa in fact has a limited degree of symmetry: both direct and second objects trigger object agreement on the verb, when only one is present (Rhodes (1990)). And so in fact does English. For most speakers, if a recipient appears as a full NP in a double object construction, it is the sole candidate for passivization; the second object is excluded:[17]

(86)   a.   Paul was handed the shovel (by Susan)
        b.   %The shovel was handed Paul (by Susan)

We can normally only passivize the theme if the recipient is expressed as an (oblique) *to*-object, as in *The shovel was handed to Paul (by Susan)*. Thus we seem to have an asymmetric construction with the recipient as direct object. But if the recipient is a pronoun, it seems to be possible to passivize the theme, at least in some dialects:

(87)   a.   No explanation was given them
        b.   The job was offered him
        c.   Fake documents were given him

Oehrle (1976:177) finds such examples scattered throughout English writing and broadcasting (the postverbal dative is almost always a pronoun).

In these cases the degree of symmetry is limited enough so that there isn't a real problem in deciding which NP should be regarded as the direct object, but in some languages the symmetrical behaviour is far more pervasive, to the point where it seems plausible to postulate multiple direct objects.

The original example of symmetrical behaviour – and still one of the most extensive – is provided by another Bantu language, Kinyarwanda (Gary and Keenan (1977); Kimenyi (1980)).

In this language, as in Chi-Mwi:ni, there can be multiple bare NPs after the verb that look like P, but in many cases more than one of them can manifest

---

[17] The '%' indicates the variable acceptability of the (b) example. Some English speakers accept it and some do not.

the grammatical properties of P, rather than only one. Two of these P-properties
are the ability to be passivized, and to be replaced by a verbal prefix when
pronominal.[18] Below is an example with three P-like postverbal NPs, and a
variant where they are all replaced by pronominal object prefixes:

(88)   a. Umugóre   a-rá-hé-er-a   umugabo ímbwa ibíryo
          woman   she-PRES-give-APPLIC-ASP   man   dog   food
          'The woman is giving food to the dog for the man'

       b. Umugóre   a-rá-bi-yí-mu-hé-er-a
          woman   she-PRES-it-it-him-give-APPLIC-ASP
          'The woman is giving it to it for him'   Kimenyi (1980:65)

And here we see any of the three being passivized (but only one at a time):

(89)   a. Ibíryo bi-rá-hé-er-w-a   umugabo ímbwa n'ûumgóre
          food   it-PRES-give-APPLIC-PASS-ASP   man   dog   by.woman
          'The food is given to the dog for the man by the woman'

       b. ímbwa i-rá-hé-er-w-a   umugabo ibíryo n'ûumgóre
          dog   it-PRES-give-APPLIC-PASS-ASP   man   food   by.woman
          'The dog is given the food for the man by the woman'

       c. Umugabo a-rá-hé-er-w-a   ímbwa ibíryo n'ûumgóre
          man   he-PRES-give-APPLIC-PASS-ASP   dog   food   by.woman
          'The man is given food for the dog by the woman'

By contrast, there are other multiple-apparent-P constructions where not all
of the bare postverbal NPs can show the grammatical object properties. For
example, a locative argument can be expressed as a bare NP after the verb
(which has a locative marker suffixed to it), along with the patient, but it is the
locative not the patient that shows the object properties of pronominalization
and passivization:

(90)   a. Úmwáalimu y-oohere-jé-ho   iishuûri igitabo
          teacher   he-send-ASP-to   school   book
          'The teacher sent the book to the school'

       b. Úmwáalimu y-a-ry-oohere-jé-ho   igitabo
          teacher   he-PAST-it-send-ASP-to book
          'The teacher sent the book to it'

---

[18] In Chi-Mwi:ini, the verbal object-marking prefixes serve as agreement markers which can co-
occur with full NPs, which can also be omitted, while in Kinyarwanda they are mutually exclusive
with NPs. See Bresnan and Mchombo (1987) for discussion of this typological difference.

    c. Iishuûri ry-oohere-j-w-é-ho   igitabo n'úúmwáalímu
       school it-send-ASP-PASS-ASP-to book   by.teacher
       'The school was sent the book by the teacher'

    d. *Úmwáalímu cy-oohere-je-é-ho   ishuûri
       teacher   he-PAST-it-send-ASP-to school
       'The teacher sent it to school'

    e. *Igitabo cy-oohere-j-w-é-ho   ishuûri n'úúmwáalímu
       book   it-send-ASP-PASS-ASP-to school by.teacher
       'The book was sent to school by the teacher'

<div align="right">Kimenyi (1980: 94–5)</div>

This shows that in the benefactive-dative-patient constructions of (89), it is reasonable to regard all the postverbal NPs as being 'direct objects', but in the locative-patient constructions of (90), only the locative.

    Crucial to the idea of multiple objects is that more than one NP be able to exhibit an object property at the same time; in Kinyarwanda this has been demonstrated only for object-pronominalization, but Bresnan and Moshi (1990) illustrate this for various other combinations of properties in the Bantu language Kichaga.

    Symmetric languages afford the problem that because object properties are shared between multiple NPs, there doesn't appear to be a clear basis for picking out a unique NP as direct object. However, we've seen that asymmetric languages can show a limited amount of symmetric behaviour, and the reverse turns out to be the case as well: Dryer (1983) shows that in Kinyarwanda there are differences between the grammatical behaviour of recipient, benefactive, and theme/patient objects.[19] Symmetry and asymmetry thus appear to be matters of degree, and a final complexity is what can be called 'split objecthood': here in a double object construction, one of the objects takes some of the P-properties, while the other takes some of the others. Dryer (1986:829–30) discusses some cases of split objectivity in Southern Tiwa, Mohawk, and other languages.

    Passivization and cross-referencing are the most widely available tests for direct-objecthood, although a wide range of other phenomena can provide evidence in particular languages. A final point is that although second objects usually appear only in the presence of direct objects, this isn't always the case; in Ojibwa for example (Rhodes (1990)), there appear to be 'secondary objects' that can appear either with or without the presence of an ordinary direct object.

3.2.2   Indirect objects   In the 'double object' constructions discussed above, there are two non-subject NPs that are similar in appearance, which may or may

---

[19] Assuming the framework of Relational Grammar, Dryer interprets the facts as evidence for a direct object / indirect object distinction, although in other frameworks there are different possibilities.

not be similar in behaviour. Another option is for the two NPs to look different. Of course one way for two non-subject arguments to look different is for one of them to be a core argument and the other an oblique; this is what happens in English, in examples like these:

(91)   a. Mary presented a watch to Tom
          b. Mary presented Tom with a gold watch

Here the arguments introduced by the prepositions *with* and *to* are classed as oblique, because of their difference in appearance and behaviour from A, S and P, which are bare NPs, and their similarities in appearance and behaviour to other non-core roles, such as instrumental and locative adjuncts. But it is also possible for the different-looking argument to present the behaviour of a core argument rather than an oblique.

This happens, for example, in Warlpiri. In Warlpiri, verbs of giving and related notions take their agent in the ergative case, their theme in the absolutive, and their recipient or related role, such as loser, in a dative. These datives are cross-referenced on the auxiliary by the ordinary object markers except for the third person singular, which is cross-referenced by *-rla*:

(92)
a. Nyuntulu-rlu ngaju-ku ka-npa-ju   karli-patu
   you-ERG   me-DAT   PRES-2SG(SUBJ)-1SG(OBJ) boomerang(ABS)-PAUCAL
   yi-nyi
   give-NONPAST
   'You are giving me a few boomerangs'

b. Ngajulu-rlu kapi-rna-rla   karli-patu
   I-ERG   FUT-1SG(SUBJ)-3(DAT)   boomerang(ABS)-PAUCAL
   punta-mi   kurdu-ku
   take away-NONPAST   child-DAT
   'I will take the boomerang / the few boomerangs away from the child'

Note in particular that a plural third person P of a transitive verb would be cross-referenced with *-jana*, on the auxiliary, while here all we have is cross-referencing of the recipient with *-ju* (92a) and *-rla* (92b).

If we assume that case should directly reflect grammatical relations when this is possible, we would want to analyse these examples by treating the theme as a direct object, and the recipient/loser as a new grammatical relation, which we can call 'indirect object' (defined as the grammatical relation, if there is one, normally associated with recipients).

But the evidence for an indirect object grammatical relation is quite equivocal. The cross-referencing on the auxiliary treats the dative NP almost exactly as if it were the direct object, showing agreement with it rather than with the

absolutive. The only difference between the cross-referencing of the recipient in a ditransitive and that of an ordinary absolutive direct object is that, in the former case, there is overt cross-referencing expressed by a morpheme *-rla*, rather than null cross-referencing. This would be straightforwardly explained if the datives were the direct objects (participating in agreement), and the absolutives were second objects (failing to agree), with the appearance of *-rla* being attributed to the case marking.

The syntactic behaviour likewise speaks against indirect object status, rather than in favour of it. The evidence comes from the use of the nonfinite complementizer *kurra*, already illustrated in (55) above, with three-argument verbs. In such cases, it seems to be more natural to interpret the subject of the *kurra*-marked verb as being the dative rather than the absolutive:

(93)   a.   Karnta-gku   ka-ju   kurdu   milki-yirra-rni
            woman-ERG   PRES-1SG(OBJ)   child(ABS)   show-put-NONPAST
            nguna-nja-kurra(-ku)
            lie-INF-OBJ.COMP(-DAT)
            'The woman is showing the child to me while I am lying down'

       b.   ??Yu-ngu-rna-rla   kurdu   parrja-rla
            give-PAST-1SG(SUBJ)-3(DAT)   child(ABS)   coolamon-LOC
            nguna-nja-kurra   yali-ki
            sleep-INF-OBJ.COMP   that-DAT
            'I gave the child sleeping in the coolamon to that one'
                                                    Simpson (1991:341–2)

Simpson (citing communications from Mary Laughren) reports that the (b) example, with the absolutive controlling the *kurra*-verb, is questionable, and that speakers prefer an interpretation where it is the recipient that's sleeping in the coolamon rather than the theme. This suggests that the dative is indeed the direct object, rather than an indirect object.

In Romance languages, on the other hand, NPs marked with the preposition *a* often have certain properties such as the ability to be cross-referenced, indicating that they are core arguments (Alsina (1996b:150–60)), but do not undergo passivization (as ordinary direct objects do), indicating that they might have a different grammatical relation, which could then be appropriately called 'indirect object' (however, Alsina (1996a:13, 150)) rejects this kind of analysis, taking the *a*-marked recipients to be simply objects, with their differences from other objects, such as the non-applicability of passivization, being due to their dative case marking).

In English, Bantu, and many other languages, on the other hand, we do not seem to find even *prima facie* plausible candidates for an indirect object grammatical relation. In these languages recipients are expressed either as

direct objects, usually in a double object construction, or as obliques. For example, the *to*-object construction in English gives no evidence of being anything other than an ordinary oblique prepositional phrase. There is no reason to set up a special indirect object relation borne by it but not by other kinds of PP.

The status of the notion of 'indirect object' is thus problematic and difficult to sort out. The top priority is to work out what properties recipients and themes do and do not share with P arguments of PTVs.

3.2.3   Other core relations   Aside from subject, object and perhaps indirect object, various other core grammatical relations sometimes seem to be motivated. An example of an unusual core grammatical relation is provided by Warlpiri. Any Warlpiri verb may be supplemented by what Hale (1973) calls an 'adjunct dative', but which we will call a 'supplementary dative', to avoid confusion with the terminology of this chapter. A supplementary dative is a dative which expresses various semantic roles, but is cross-referenced as an indirect object. If associated with a verb with no special marking, the supplementary dative is interpreted as a beneficiary:

(94)
a. Ngarrka-ngku ka-rla   kurdu-ku karli   ngurrjuma-ni
   man-ERG   PRES-3(DAT) child-DAT boomerang(ABS) make-NONPAST
   'The man is making a boomerang for the child'

b. Ngarrka-ngku ka-rla-jinta   kurdu-ku miyi   karnta-ku
   man-ERG   PRES-3(DAT)-3(DAT) child-DAT food(ABS) woman-DAT
   yi-nyi
   give-NONPAST
   'The man is giving food to the child for the woman'
   or: 'The man is giving food to the woman for the child'

Example (94b) shows that the supplementary dative can co-occur with an indirect object, and is thus a distinct grammatical relation. *Jinta* is the form assumed by the second of two cross-reference markers both referring to a third person singular dative (Hale (1973:336)).

The interpretation of the supplementary dative may be altered by adding to the verb one of a number of so-called 'preverbs' (which have a variety of additional functions in Warlpiri). Thus, with the preverb *marlaja*, the adjunct dative indicates the entity who brings about the situation described by the sentence. With the preverb *piki(-piki)*, the dative represents an entity from which some participant is in danger:

(95)

a. Kurdu-ngku ka   miyi   nga-rni
   child-ERG   PRES food(ABS) eat-NONPAST
   'The child is eating food'

b. Kurdu-ngku ka-rla   karnta-ku   miyi   marlaja-nga-rni
   child-ERG   PRES-3(DAT) woman-DAT food(ABS) CAUSE-eat-NONPAST
   'The woman brought about the circumstance that the child is eating food'

(96)

a. Ngarrka-ngku ka   yujuku   nganti-rni
   man-ERG   PRES humpy(ABS) build-NONPAST
   'The man is building a humpy [bush shelter]'

b. Ngarrka-ngku ka-rla   warlu-ku piki-nganti-rni
   man-ERG   PRES-3(DAT) fire-DAT DANGER-build-NONPAST
   yujuku
   humpy(ABS)
   'The man is building a humpy in danger of fire [either man or humpy
   is in danger]'

Since the semantic role of the supplementary dative is determined by the form
of the verb, its status as a core grammatical relation is confirmed.

Supplementary datives are probably best viewed as the results of a lexical
process which derives from one lexical item another with an additional argument
whose semantic role is determined by which preverb, if any, is added.

It is quite common for benefactives to be added by a lexical operation of this
sort, although the technique employed in Warlpiri is unusual. More commonly,
the benefactive takes on the appearance and at least some of the properties
of a direct object. In English, for example, we can express a benefactive as a
for-adjunct or as what looks like a direct (first) object:

(97)   a. Bruce barbecued the steak for Darlene
         b. Bruce barbecued Darlene the steak

Although the benefactive object in (97b) looks like a direct object, it is
behaviourally somewhat different, since for most speakers it cannot passivize:
Darlene was barbecued the steak by Bruce (Fillmore (1965)). It is not clear
whether we should think of benefactive objects as having a different grammat-
ical relation from ordinary direct objects, or whether the differences are simply
a consequence of the semantic role of the benefactives.

In many Bantu languages, benefactives can only be expressed as derived
object-like NPs, benefactive adjuncts being absent. Furthermore, the benefactive
objects take on object properties more readily than in English, being freely
cross-referenced, passivized, etc. Similar processes also add arguments with a

wide range of other semantic roles, such as instrument, locative, reason, etc. Such processes are widely discussed under the title of rules of 'applicative' formation (Baker (1988); P. K. Austin and Bresnan (1997)).

3.3   Syntactic ergativity

In many languages with an ergative case-marking system, such as Warlpiri, the syntax appears to be organized along subject–object lines, as originally argued by Anderson (1976), and confirmed by much subsequent work, as reviewed and extended in Simpson (1991). But there are also languages in which at least some of the syntax is organized along absolutive–ergative lines, with rules targeting P/s rather than A/s. This phenomenon is called 'syntactic ergativity'. Languages that appear to be overwhelmingly ergative in their syntax are quite rare (there is only one well-described example, Dyirbal (Dixon (1972))); for languages with syntactic ergativity, the usual case is for some subject-like properties to apply to the P/s, others to the A/s, a situation that is called 'mixed ergativity'.

In this subsection we will introduce syntactic ergativity in the Australian language Yidiɲ (Dixon (1977a)), and then provide some discussion of the more extensively ergative (and therefore more unusual) language Dyirbal. These two languages are concisely described and compared in Dixon (1977c). Then in the next section we will consider mixed ergativity together with another problematic kind of system of grammatical relations, the 'Philippine type', and will use these to motivate some revisions to our conception of grammatical relations.

Yidiɲ, like Warlpiri and most other Australian languages, has rather free word order (though there are strong preferences), relying entirely on NP marking to code syntactic functions. Under certain circumstances, the components of an NP may be split (Dixon (1977a:268–71)), but this is far more restricted than in Warlpiri (or Dyirbal, which is similar to Warlpiri in having very free word order). The NP marking system is of the split ergative type, with different categories of nominals having different systems of case forms.

The three relevant categories are common nominals (nouns and adjectives), pronouns, and deictics. Common nominals inflect ergatively, taking an ergative form in A function, an absolutive form in P/s function. Pronouns (existing only for first and second persons; for third person reference demonstratives are used) take an accusative in P function and a nominative in A/s function. Deictics (comprising demonstrative and interrogative/indefinite pronouns, the former also serving as third person pronouns) have two stems, human and nonhuman. Humans may only be referred to by a human stem, while nonhumans may be referred to by either, the use of the human stem being more likely the more humanlike the referent of the NP. Human deictics take an ergative A form, an accusative P form, and an absolutive s form, while nonhuman

deictics inflect like common nominals, except that for some there is an optional accusative.

Some examples illustrating case marking for personal pronouns and common nouns are the following:

(98)   a. ŋayu   maŋga:-ɲ
           I(NOM)   laugh-PAST
           'I laughed'

       b. buɲa   maŋga:-ɲ
           woman(ABS)   laugh-PAST
           'The woman laughed'

       c. ŋaɲa-ɲ buɲa:-ŋ   wuɹa:-ɲ
           I-ACC   woman-ERG   slap-PAST
           'The woman slapped me'

       d. ŋayu   buɲa   wuɹa:-ɲ
           I(NOM)   woman(ABS)   slap-PAST
           'I slapped the woman'

       e. Waguɖa-ŋgu guda:ga   wawa:-l
           man-ERG   dog(ABS)   see-PAST
           'The man saw the dog'

The evidence for syntactic ergativity in Yidiɲ comes from the subordinate clause constructions of the language. These are similar in function to the relative clauses of Diyari – see (67), (68) – having what from the English point of view are a variety of relative and adverbial interpretations. There are four morphological types of subordinate clauses, 'dative', 'causal', 'purposive' and 'apprehensional', each with a different ending on the subordinate verb. The first three types are quite similar in their behaviour, while the apprehensional clauses are somewhat different.

There is no switch-reference system in Yidiɲ. But there is in the dative, purposive and causal subordinate clauses a near requirement that if the matrix and subordinate clauses contain coreferential NPs (about 85 per cent do in Dixon's texts (Dixon (1977a:323))), this NP should have P/s function in both clauses. This requirement is absolute for clauses with a relative interpretation, that is, for those in which the coreferentiality is essential to the function of the clause, though it is occasionally violated by those with adverbial interpretations (Dixon (1977a:323–49)). Furthermore, the coreferential NP in the subordinate clause may only be ellipsed if it is in P/s function (Dixon (1977a:332–3)).

Thus we can use the dative subordinate clause construction DATSUB, signaled by the verbal suffix -*ɲunda*, which expresses simultaneous action, to combine (98a) and (98c) to yield either (a) or (b) below:

(99) a. ŋayu   manga:-ɲ   (ŋaɲa-ɲ) buɲa:-n   wuɹa:-ɲunda
          I(NOM)   laugh-PAST   I-ACC   woman-ERG   slap-DATSUB
          'I, who was slapped by the woman, laughed'

      b. ŋaɲa-ɲ   buɲa:-ŋ   wuɹa-ɲ   / (ŋayu)   manga-ɲunda
          I-ACC   woman-ERG   slap-PAST   I(NOM)   laugh-DATSUB
          'I, who was laughing, was slapped by the woman'

In (99a) the matrix coreferential NP is s, and the subordinate is P; in (99b), the matrix coreferential NP is P and the subordinate one is s. s-s and P-P combinations are also possible. In these examples the matrix and subordinate coreferential NPs differ in their case form, since the shared NP is a personal pronoun, and therefore has a nominative form in s function and accusative in P. But the same coreference possibilities would exist if the NPs in both clauses were common nominals, with the same case forms in both clauses. The examples of (99) also illustrate optional omission of the subordinate clause NP: it is also possible to omit the main clause NP, or, rarely, both.

   If one of the coreferential NPs is A, the clauses cannot normally be combined as they are. Rather, a rule that is both similar to and different from the passive of languages like English must be used to convert the A to s function, except, very rarely, when the clause is adverbial in sense, and the coreferentiality is 'accidental' (not essential to the function of the clause).

   All transitive verbs have a so-called 'antipassive' form, derived by adding the suffix designated -ḍi-n by Dixon (the *n* represents the conjugation class of the antipassivized verb, which manifests itself by its effects on the form of what follows it). The role normally expressed as A is then expressed as s, while the role normally expressed by P is expressed by an NP in the dative or locative case, the choice determined by humanness in the same way as the choice between human and nonhuman deictic stems: NPs referring to humans must be dative, while those referring to nonhumans may be either dative or locative, but are more likely to be dative the more like humans they are (Dixon (1977a:110–12)). This alternation extends to many but not all values of the dative and locative case forms.

   The antipassive construction is illustrated below, where (100a) is the antipassive of (98e), and (100b) is the antipassive of (98d):

(100) a. wagu:ḍa   gudaga-nda/-la   wawa:-ḍi-ɲu
           man(ABS)   dog-DAT/LOC   see-ANTIPASS-PAST
           'The man saw the dog'

       b. ŋayu   buɲa:-nda   wuɹa:-ḍi-ɲu
           I(NOM)   woman-DAT   slap-ANTIPASS-PAST
           'I slapped the woman'

In the change from (98e) to (100a), the agent changes its case from ergative to absolute, since it is a common noun, but the pronominal agent in (98d) (100b) has no case change, since it takes the nominative form for both A and S functions.

Antipassives appear to be virtually exact paraphrases of the corresponding non-antipassive constructions. Questions, for example, will often be answered in the antipassive simply for the sake of injecting grammatical variation into the discourse (Dixon (1977a:118)). Like passives in English, however, antipassives appear to be secondary constructions in that they have greater morphological complexity in the verb, and are not used without some reason (one might answer a question in the antipassive to vary the style, but wouldn't ask it that way out of the blue).

The antipassive permits us to get the effect of combining (98a) and (98d), in which the shared NP is S in one clause and A in the other. Example (98d) is converted to its antipassive form (100b), and we get the sentences below as the result (following Dixon's presentation, '/' serves as a clause-separator):

(101) a. ŋayu   maŋga:-ɲ / (ŋayu) buɲa:-nda   wuɹa:-ḍi-ɲunda
          I(NOM) laugh-PAST I(NOM) woman-DAT slap-ANTIPASS-DATSUB
          I, who was slapping the woman, laughed; I laughed while slapping
          the woman

       b. ŋayu   buɲa:-nda   wuɹa:-ḍi-ɲu   / (ŋayu) maŋga-ɲunda
          I(NOM) woman-DAT slap-ANTIPASS-PAST I(NOM) laugh-DATSUB
          I, who was laughing, slapped the woman; I slapped the woman
          while laughing

The purposive and causal subordinate clauses, which I will not illustrate here, behave in exactly the same way. In these three types, we have a strong preference (though there are a few counterexamples) for a shared NP to be in P/S function in both the subordinate and matrix constructions. Furthermore this requirement must be met if the clause is to be interpreted as an NP-modifier or if the subordinate clause instance of the NP is to be deleted (Dixon doesn't state whether deletion of the matrix NP obeys this condition). The syntactic rather than semantic character of the principles constraining clause combination is revealed by the fact that the antipassive, which turns an A into an S, permits an agent NP to come to satisfy them.

These principles treat P and S equivalently, and therefore motivate establishing a grammatical relation, which we shall call 'absolutive', expressing P and S functions. For A function we would propose another grammatical relation, 'ergative'. In Yidiɲ there is very little further corroboration for this analysis.

But in Yidiɲ's southerly neighbour Dyirbal, the case for P-S identification is much stronger. All of the complex sentence constructions of the language (two to four, depending on how one counts) provide evidence for treating

P and S as having one grammatical relation, and there are various morphological phenomena that do as well.

Symptomatic of the difference between the two languages are the differences in their sentential coordination constructions. One of the most characteristic features of Dyirbal discourse is that long sequences of coordinate clauses tend to be strung together in a 'topic chain', in which all the conjuncts have a shared NP in P/S function, that is, with the absolutive grammatical relation (Dixon (1972:130–2)).

In Yidiɲ on the other hand, one does not find such topic chains: coordinations (expressing simultaneous action of two or three clauses containing a shared NP) are reasonably common, but not the sequences of up to a dozen or more clauses that one finds in Dyirbal (Dixon (1977a:388)). Furthermore the shared NP is not always constrained to have the absolutive grammatical relation. Rather, if it is a common nominal, it must be in the absolutive case in both clauses, while, if it is a pronoun, it must be in the nominative in them (Dixon, 1977a:388–92)). Thus, if the shared NP is a pronoun, it will have A/S function in both clauses, but if it is a noun, P/S.

We thus have a *prima facie* case, strong in Dyirbal, but weaker in Yidiɲ, that these languages have an essentially different sort of syntactic organization from that found in standard 'subject-oriented' systems such as English. They seem to lack a 'subject' grammatical relation (under the definition presented in this chapter) but have instead an 'absolutive' grammatical relation expressing P/S function. This raises the question of exactly what kind of a grammatical relation this 'absolutive' is. Is it simply the familiar 'subject', with a different alignment to semantic roles, or something more essentially distinct, suggesting a change in our conception of of how grammatical relations work? A problem with the former view is that there are a substantial number of languages for which it is unclear whether they are syntactically ergative or not,[20] which is troubling because one would not expect frequent ambiguity about a basic feature of a language's organization. In the next section we consider evidence that the latter view is in fact the case.

In the next section we will consider a variety of phenomena that motivate a reconsideration of grammatical relations.

4   Reconsidering grammatical relations

In the conception of grammatical relations that has been assumed by our work so far, each NP in a clause bears a single grammatical relation. One problem for this view arises from syntactic ergativity: in some languages with

---

[20] Dixon (1977a:393) observes that participants in a conference session devoted to whether various Australian languages were syntactically ergative or not were frequently doubtful of the correct treatment of their languages, often changing their minds in the course of preparing final versions of their papers.

ergative features, it is unclear whether the subject grammatical relation should be regarded as expressing A/S or P/S functions, since the evidence is weak, or, as we shall see below, contradictory. But syntactic ergativity is not the only problem for grammatical relations. Another set of difficulties comes from the so-called 'Philippine type' of language structure, which seems in a sense to have two systems of grammatical relations functioning at the same time. In this section we will suggest a solution to both problems, which originated with some proposals by Keenan (1976c) about the nature of the subject concept, and has been developed by many other authors since then.[21]

The basic idea of the solution is that the familar concept of subject should in fact be split into two concepts, one associated with the semantic role of agent, the other with the pragmatic role of topic. In English these two concepts pick out the same NP in the sentence, but in certain others, such as syntactically ergative languages and the Philippine type, they don't. Therefore in English we have the grammatical relation of subject, while in some other languages we must distinguish what we might call 'a-subject' (agent-oriented) from 'p-subject' (pragmatic, topic-oriented, pivot).

We will first show how this idea helps with the problem of 'mixed ergativity', where a language shows a combination of ergative–absolutive and nominative–accusative organization. We will then show how it applies to the problematic features of the Philippine type.

4.1   *Mixed syntactic ergativity*

Early work on syntactic ergativity assumed that languages would either show ergative–absolutive or nominative–accusative organization, depending on whether their sentence structures treated S and A alike (the majority), or S and and P. But this assumption proved initially to be dubious and, ultimately, false.

An initial reason for doubt is that the evidence for setting up the grammatical relations one way or another in ergative languages is often rather weak, as we saw for example in Yidiɲ. More serious is the fact that there are often contradictory indications about which way they should be set up. Sometimes one can make a case that the syntactic phenomena are showing ergative–absolutive organization, and that the apparently nominative–accusative phenomena are being semantically conditioned, but there are also instances where it is clear that some syntactic phenomena are ergative–absolutive, while others are nominative–accusative. An example of the first type is Yidiɲ (Dixon (1977a)), of the second, Inuit (Bittner (1994); Manning (1996)). Yidiɲ also illustrates the rather common case where the available evidence about grammatical relations is rather scanty, so we will discuss it first.

---

[21] Such as Schachter (1976, 1977), Foley and Van Valin (1984), Guilfoyle, Hung, and Travis (1992), Kroeger (1993), Wechsler and Arka (1998), and others writing on Austronesian languages, and Dixon (1979), Bittner (1994), and Manning (1996) on syntactic ergativity.

We have already seen the evidence for an s/p grouping (putative p-subject) in Yidiɲ (see (98)). But there are two constructions that treat A and s alike as opposed to P. The first is the imperative: imperatives require a second (or occasionally a first) person pronoun in s/A function (Dixon (1977a:370–1)). Imperatives with s and A addressees are illustrated below:

(102)   a. (ɲundu) guwa gali-n
           vou(SG)   west go- IMPER
           '(You) go west!'

        b. (ɲundu:ba) buɲa   wawa
           you(PL)   woman watch(IMPER)
           '(All of you) watch the woman!'

The second involves a number of particles whose grammar treats s and A alike (Dixon (1977a:372–82, 387)). For example the particle *gana:ŋgar* indicates that the referent of the NP in s or A function was the first to perform a certain action:

(103)   a. ŋayu   gana:ŋgar gali:-ɲ
           I(NOM)   first   go-PAST
           'I went first'

        b. ŋayu   gana:ŋgar gunda:-l
           I(NOM)   first   cut-PAST
           'I was the first person to cut [that tree]'

But these phenomena don't constitute a truly compelling case for saying that the syntax is recognizing an s/A category, because what might be happening is that the phenomena have a semantic rather than a syntactic basis. In particular, the semantics of both the imperative and the particle constructions might be such that they involve an agentive argument in their interpretation, and impose constraints on it. Since Ps are never agents, Ps won't be able to be involved in these constructions, for reasons that are quite independent of how the syntax is organized.

In principle, one could investigate this issue by looking at intransitives with nonagentive s, and also antipassives, but Dixon doesn't provide significant discussion of this, although he does provide a suggestive example of the 'cessation' particle *wala* modifying the presumably nonagentive s of *die*:

(104)   ŋayu wala wula:ɲ ŋayu galwayala burgiŋ
        I   finish die   /I   spirit   walk about
        'I really did die; I'm walking about as a spirit now'

                                                      Dixon (1977a: 375)

Regardless of the uncertainties of Yidiɲ (which are typical of what one finds with data from field work), there are languages with clearer cases of mixed ergativity, such as Inuit, which we now consider.

Inuit, also known as Central Canadian Eskimo, is a language with relatively free word order, an extremely rich system of word-formation, and a system of case marking and cross-referencing on verbs that is somewhat reminiscent of Australian Aboriginal languages, as well as older or conservative Indo-European ones such as Sanskrit or Russian. The case marking is ergative, with the ergative case being identical to the possessive; this ergative/possessive case is traditionally called the 'relative'. Intransitive verbs agree with their s, transitives with A and P, via a complex system of cross-referencing affixes, which cannot convincingly be resolved into distinct A and P markers. Basic case marking and cross-referencing are illustrated in these examples:

(105)   a. Atuagaq   ataasiq   tikis-sima-nngi-la-q
           book(ABS) one(ABS) come- PERF-NEG-INDIC-3SG
           One book hasn't come yet

        b. Juuna-p   atuagaq   ataasiq   tigu-sima-nngi-la-a
           Juuna-ERG book(ABS) one(ABS) get-PERF-NEG-INDIC.3SG-3SG
           There is a book which Juuna hasn't got yet

Note that the absolutive case is 'marked' by the absence of any case-affix. So the case marking is ergative, but the morphology of the verbal cross-referencing is too complex to support a clear judgement of whether it is ergative/absolutive or nominative/accusative. However, there are a number of phenomena showing syntactic ergativity in Inuit, collected by various researches over the years, and summarized and discussed by Manning (1996:83–191). Here I will present two: participial relative clauses, and the 'wide scope' that applies to the absolutive.

The latter effect, discovered by Bittner (1994), is illustrated by the somewhat peculiar translations given to the examples above: the absolutive is interpreted as something that exists, about which the negative assertions are made, such as that it hasn't come, or that Juuna doesn't have it. The sentences do not have the following as glosses, where the existence of books is not assumed:

(106)   a. No books have come yet
        b. Juuna hasn't got any books yet

In Inuit, an absolutive argument will thus have semantically 'wide scope' over negative markers in the verbal morphology. An ergative-marked A on the other hand can have either wide scope over, or narrow scope under, a negative:

(107)   Atuartu-p   ataatsi-p Juuna   uqaluqatigi-sima-nngi-la-a
        student-ERG one-ERG Juuna(ABS) talk.to-PERF-NEG-IND-3SG-3SG
        'No student has talked to Juuna yet'
        'One student hasn't talked to Juuna yet'

Wide scope is a property often (but not necessarily) associated with subjects, so not only is the absolutive showing a distinctive property, but also one that is subject-like.

Our other example showing syntactic ergativity is relativization. Inuit has a series of participial moods that can be used to form relative clauses, but only relativizing on s or p in the relative clause (or, marginally, their possessors):

(108)   a. Miraaq   kamat-tu-q
            child(ABS) angry-REL.INTRANS-SG
            'the child that is angry'

        b. Nanuq   Piita-p   tuqu-ta-a
            polar.bear(ABS) Piita-ERG   kill-TR.PTCPL-3SG
            'a polar bear killed by Piita'

        c. *Angut   aallaat   tigu-sima-sa-a
            man(ABS)   gun(ABS)   take-PERF-REL.TR-3SG.3SG
            intended: 'the man who took the gun'   Bittner (1994:56–7)

So we see that (a), with relativization on s, and (b), with relativization on p, are acceptable, while (c) with attempted relativization on a is not.

Participial relativization and wide scope are properties of s/p as opposed to a. They are also related to definiteness and topicality, and so characteristic of the properties of pivots in Tagalog. It is therefore natural to classify them as p-subjects.

There are also phenomena which involve s/a but not p, where this restriction can't be explained away as a simple consequence of the meaning. The one we will discuss here involves the extremely complex verb-formation system of the language – for more details see Manning (1996:101–47).

Inuit is famous for a system of word-formation whereby more complex verb forms can be derived from simpler ones by suffixing formatives that are called 'post-bases' (whether they are affixes or not is controversial). These complex verb forms take on functions achieved by auxiliaries and complement structures in English. One of these suffixes means 'want', and it attributes the desire to s/a rather than p:

(109)   a. Hansi   sinik-kuma-vuq
            Hansi(ABS) sleep-want-INDIC.INTRANS.3SG
            'Hansi wants to sleep'

        b. Aani-p   miiqqat   ikiur-uma-v-a-i
            Aani-ERG   children(ABS) help-want-INDIC-TR-3SG.3PL
            'Aani wants to help the children'

In particular in the transitive (b) sentence, the desirer is Aani, the A of help, rather than the children, the P. One might suspect that this effect is caused by the semantics of the post-base, but there is clear evidence that it isn't: one can ascribe the desire to the helpee by passivizing the *ikiur* 'help', and attaching 'want' to the result:

(110)   Miiqqat   Aani-mit   ikiur-niqar-uma-pp-u-t
         children(ABS) Aani-ABL   help-PASS-want-INDIC-INTRANS-3SG
         'The children want to be helped by Aani'

In the passive, the former A is expressed as an ablative oblique, while the former P is expressed as an absolutive, and is evidently now an S. And concomitantly, it is interpreted as the desirer. This shows that which argument is understood as the desirer is determined by the grammatical structure rather than the semantic roles.

We can accommodate this mixture of ergative and non-ergative features by splitting the subject grammatical relation into two distinct and overlapping ones, 'p-subject' identified with P/s function, and 'a-subject' identified with A/s function. The phenomena showing syntactic ergativity are sensitive to p-subject, while the ones treating A and S alike are sensitive to a-subject. How do unmixed syntactically ergative languages such as Dyirbal fit into this picture? Clearly they have p-subjects following the same principle with mixed ergative languages, but there are two possibilities for a-subject. The first is that it is present, but the evidence for its existence has not yet been found and reported, the second is that in these languages a-subject does not exist. The issue will be discussed further below.

We now proceed to extend the split subject hypothesis to apply it to the Philippine type.

4.2   *The Philippine type*

The current literature on grammatical relations and the Philippine type essentially begins with the analysis and discussion of Tagalog in Schachter (1976, 1977), which itself grew in part out of the discussion of the concept of 'subject' in Keenan (1976c), as well as previous Philippinist literature. A distinctive feature of these languages is the possession of what has often been called a 'focus' system, in which one NP is singled out for special treatment in a manner reminiscent of subjects in more familiar languages, but with sufficiently different behaviour to have made it controversial whether the singled-out NP should indeed be seen as a subject. We begin with a brief account of Tagalog, including the focus system, and then consider the issues that it raises for the notion of subject, and also the analysis of syntactically ergative languages.

Tagalog has verb-initial order, with NPs appearing in free order after the verb, with their functions marked by prepositional NP-markers (there is also a topicalization construction (Kroeger (1993:43–4,123–4)), in which any NP may be placed in front of the verb). Verbs are traditionally considered as taking three types of 'core' arguments, labelled 'Actor', 'Object', and 'Directional' by Schachter and Otanes (1972).

Actor and object are marked by *ng* (pronounced [nəŋ]) which I will gloss as ACT when it marks an actor, OBJ when it marks an object, using two glosses rather than one in order to make the examples easier to follow. Directionals are marked by *sa*, unless they are 'pivot', as will be discussed below. The traditional names for these types of argument are semantically suggestive but not fully accurate: actors needn't be agents, and directionals needn't be (semantically) directional. There are also various sorts of adjuncts: benefactives, outer locatives, instrumentals, etc.

One of the arguments (or, more rarely, one of the adjuncts) must be chosen to be what we here call the 'pivot' (the terms 'focus' and 'topic' are also sometimes used), which we will later identify as the p-subject. The pivot bears the marker *ang*, glossed PIV, instead of the marker that would otherwise appear, and is obligatorily understood as being definite. The type of argument or adjunct that is chosen as the pivot is indicated by affixes on the verb. Below is illustrated an array of pivot choices for the verb *alis* 'take out', which has actor, object, and directional arguments, and here appears with a benefactive adjunct as well (AP = actor pivot, OP = object pivot, DP = directional pivot, BP = benefactive pivot):

(111) a. Mag-a-alis   ang babae   ng bigas sa sako para sa bata
           AP-FUT-take.out PIV woman OBJ rice   DIR sack BEN   child
           'The woman will take some rice out of a/the sack for a/the child'

       b. A-alis-in   ng babae   ang bigas sa sako para sa bata
           FUT-take.out-OP ACT woman PIV rice   DIR sack BEN   child
           'A/the woman will take the rice out of a/the sack for a/the child'

       c. A-alis-an   ng babac   ng bigas ang sako para sa bata
           FUT-take.out-DP ACT woman OBJ rice   PIV sack BEN   child
           'A/the woman will take some rice out of the sack for a/the child'

       d. Ipag-a-alis   ng babae   ng bigas sa sako ang bata
           BP-FUT-take.out ACT woman OBJ rice   DIR sack PIV child
           'A/the woman will take some rice out of a/the sack for the child'

In these examples, the choice of determiners in the glosses is significant, and is governed by two principles, the one already mentioned that the pivot is always understood as definite, and another to the effect that non-pivot objects

(but not actors or directionals) are normally understood as indefinite. This is an indication that the pivot is associated with a topic-like pragmatic concept.

The first analyses of these constructions treated the *ang*-phrase as an ordinary subject, and the non-AP (actor-pivot) forms essentially as passives (see Bloomfield (1917) and other treatments discussed by Kroeger (1993:19)), and such analyses have also been proposed by generative authors such as Schwartz (1976) and Bell (1976) for other Philippine languages. The AP form (111a) is taken as the primary form with the actor as subject, the others as passives, with the actor 'demoted' from the subject relation, and some other NP serving as subject.

That the pivot bears a subject-like grammatical relation is made clear by the fact that it is targetted by certain principles which tend to target subjects in various languages. Schachter (1976, 1977) presents three of these: relativization, quantifier launching, and an inability to appear as something whose existence is asserted in an existential sentence. We illustrate the first two.

Tagalog relative clauses take the form of sentences with ellipsed pivot. The ellipsed pivot is understood to be the head NP that the clause is modifying. Hence to relativize on an actor, one uses an AP verb; to relativize on an object, an OP (object-pivot) verb:[22]

(112) a. Matalino   ang lalaki-ng   b[um]asa ng diyaryo
           intelligent PIV man-LINK [AP]-read OBJ newspaper
           'The man who read a newspaper is intelligent'

       b. Interesante ang diyaryo-ng   b[in]asa-Ø   ng   lalaki
           interesting PIV newspaper-LINK [PERF]-read-OP ACT man
           'The newspaper that the man read is interesting'

(113) a. *Matalino   ang lalaki-ng   b[in]asa-Ø   ang diyaryo
           intelligent PIV man-LINK [PERF]-read-OP PIV newspaper
           'The man who read a newspaper is intelligent'

       b. *Interesante ang diyaryo-ng   b[um]asa ang lalaki
           interesting PIV newspaper-LINK [AP]-read PIV man
           'The newspaper that the man read is interesting'

The -*ng* suffix in these examples, glossed LINK, is an element often called a 'linker', which has various functions in the grammar: here it is regularly placed on a word in an NP immediately before a relative clause modifying that NP. In (112), we see that the pivot can be relativized upon; in (113) we see that non-pivots cannot be relativized upon. Relativization therefore targets the pivot.

---

[22] Some of the affixes are infixed; these are enclosed in square brackets, in the forms and glosses, rather than being separated from their stems by dashes.

The other pivot-targetting process is a 'Quantifier Launching' phenomenon. Tagalog quantifiers normally occur within the NP they modify, but, for some speakers, the quantifier *lahat* may also be placed in an adverbial particle position directly after the verb (Schachter and Otanes (1972:147–8)).

Such a 'floated quantifier' may modify only the pivot, not a non-pivot:

(114)   a. Ø-su-sulat   lahat ang mga bata ng mga liham
            AP-FUT-write all   PIV PL   child OBJ PL   letter
            'All the children will write letters'

        b. Su-sulat-in   lahat ng   mga bata ang mga liham
            FUT-write-OP all   ACT PL   child PIV PL   letter
            'The/some children will write all the letters'
            not 'All the children will write the letters'

The pivot thus functions as target for a number of grammatical processes, indicating that it is the bearer of a grammatical relation.

Although the phenomena of (112–14) show that the pivot has a subject-like grammatical relation, there are problems with treating the non-AP forms as passives. The OP (object-pivot) and DP (directional-pivot) forms are extremely common, rather than being relatively rare, as is typically the case with passives. Furthermore, they are not morphologically more complex than the putatively primary AP forms, but merely have different affixes, not additional ones.

But a much more serious problem with the passive analysis was first delineated by Schachter (1976, 1977), and then substantially reinforced by Kroeger (1993). Schachter observed that the actor showed a substantial number of properties that are characteristic of subjects, regardless of whether or not it was the pivot: he cited three subject properties for non-pivot actors, and three subject properties for non-actor pivots (an actor pivot would have all six). Later work shows that these claims need to be qualified substantially. For example Andrews (1985:143–4), to be reviewed just below, showed that two of the supposed subject properties did not in fact discriminate between grammatical relations in Tagalog. But on the other hand, Kroeger (1993) showed conclusively that the non-pivot actor is a core argument rather than an oblique, which is what a passivized A would be, decisively ruling against the passive analysis, and confirming Schachter's essential insight, since an oblique cannot be a core argument.

An alternative to the passive analysis which has sometimes been proposed is the ergative analysis (Gerdts (1988); T. S. Payne (1982)), in which the OP forms rather than the AP forms are taken as basic, and the AP as antipassives. But this analysis faces essentially the same difficulties as the passive analysis, but in a slightly different form: on the one hand the AP forms are too common to be plausibly regarded as antipassives, and, on the other hand, non-pivot

patients also pass the tests for being core arguments, whereas an antipassive patient is supposed to be oblique. So neither the passive nor the antipassive analysis is genuinely satisfactory, because neither of them accommodates the roughly equal status of the AP and OP constructions as basic in the language, nor the core argument status of the non-pivot A and O. Therefore we need a new analysis. We will first examine the evidence more closely, and then present a solution.

We begin by looking at the arguments originally advanced by Schachter to the effect that the actor should be regarded as a sort of subject. These arguments depend on the roles of the actor in the phenomena of reflexivization, imperative formation, and complement subject ellipsis. Although these are weaker than one would hope (especially the second one), it is worth spending some time on them because they illustrate the kinds of issues that must be dealt with when arguing for grammatical relations in a language.

The first is the observation that actors in Tagalog can be antecedents of reflexive pronouns regardless of whether they are pivots or not:

(115)　a. Nag-alala ang lolo　　　sa　kaniya-ng sarili
　　　　　　AP-worry PIV grandfather DIR his-LINK　self
　　　　　　'Grandfather worried about himself'

　　　　b. In-alala-Ø　　　ng　lolo　　　ang kaniya-ng sarili
　　　　　　PERF-worry-OP ACT grandfather PIV his-LINK　self
　　　　　　'Grandfather worried about himself'　　　Schachter (1977:292)

Schachter also shows that non-actors cannot be the antecedents of reflexive actors, so that the actor, but not the pivot, is relevant to reflexivization possibilities. This is taken to be relevant to the subject status of the actor because the ability to antecede reflexive pronouns is one of the characteristic properties of subjects listed in Keenan (1976c).

But the problem with this is that although it is usually possible for subjects to antecede reflexive pronouns, and sometimes (as in Malayalam) only possible for subjects to do so, there are also languages where non-subjects and indeed non-core arguments can antecede reflexive pronouns, such as for example English:

(116)　　John talked to Mary about himself/herself

And in Tagalog it is possible for arguments that are neither actors nor pivots to antecede reflexives:

(117)
a. In-i-abot　　　niya　　sa　bata　ang kaniya-ng sarili-ng　larawan
　　PERF-OP-hand he(ACT) DIR child PIV his–LINK　self-LINK　picture
　　'He$_i$ handed the child$_j$ a picture of himself$_{i, j}$'

b. T[um]anggap ang Rosa ng sulat para sa bata sa kaniya-ng sarili
   [AP]-receive PIV Rosa OBJ letter BEN child DIR her-LINK self
   'Rosa$_i$ received a letter for the child$_j$ from herself$_i$/him-herself$_j$'

Bell (1976:30, 157) notes essentially the same facts in the closely related language Cebuano. She suggests that Cebuano reflexivization is governed by a principle referring to semantic roles rather than grammatical relations, the Thematic Hierarchy Condition of Jackendoff (1972) (she also notes some constraints involving surface word order). The same kind of analysis seems indicated for Tagalog. Since Tagalog reflexivization, as opposed to that of Malayalam, seems to function in terms of semantic roles rather than grammatical relations, it does not provide evidence that actor is a grammatical relation independent of pivot. However, the argument does at least show that actors outrank some other NPs on a grammatically relevant hierarchy, since actors can antecede reflexives with more semantic roles than other NPs.

Next we look at imperatives. Imperative sentences have the verb in a 'base' form with focus-marking, but no aspectual marker. Schachter observes that they can have the (second person) addressee as either pivot or non-pivot, as long as it is actor:

(118) a. Mag-bigay ka   sa kaniya ng kape
          AP-give   you(PIV) DIR him   OBJ coffee

       b. Bigy-an mo   siya   ng kape
          give-DP you(ACT) him(PIV) OBJ coffee
          'Give him some coffee!'

In (a), the addressee-actor is pivot, in (b) it isn't (note that the pronouns are morphologically fused with their function markers). Both are good as imperatives. Schachter's claim is that the only actors tolerated in imperative sentences are second person pronouns (1977:291). But there are two reasons why the evidence given doesn't show that there really is an actor grammatical relation.

One reason is that the semantics of imperatives are such that one would expect them to occur with second person agents, and no syntactic phenomena have been adduced to show that the relevant notion is a grammatical relation rather than a semantic role. In fact there is evidence that imperative addressees do have an agentivity condition on them: an imperative verb cannot be an 'Involitive' form (Schachter and Otanes (1972:402)), involitives being verb forms that express accidental or involuntary action.

But there is also a deeper reason. The verb form used for imperatives is not restricted to imperative usage. It is rather used in a range of constructions expressing a desire that something happen, called 'hortatives' if the subject is first person plural, and 'optative' if the subject is first person singular:

(119)   a. Walis-an   natin   ang sahig
             Sweep-OP us(DU.ACT)   PIV floor
             'Let's us two sweep the floor'

         b. Walis-an   nila   ang sahig
             Sweep-OP they(ACT)   the floor
             'I want them to sweep the floor'

<div align="right">Schachter and Otanes (1972:407–9)</div>

There are various constraints on these constructions, and on the use of various particles with them. In (119b), for example, the subject cannot be a third person full NP. It is possible that careful analysis of these constraints could provide grounds for individuating a specific imperative construction with a second person actor, but this work has not yet been done. So the imperatives (construed as a type of speech act) provide no evidence relevant to grammatical relations in Tagalog, not only because there is no evidence of syntactic restrictions on them, but furthermore due to the absence of even a *prima facie* case that there is a distinct imperative construction in the grammar.

The final phenomenon is argument ellipsis in complement constructions (Schachter (1977:293)). Schachter argued that actors and only actors could be ellipsed, regardless of whether they were pivot:

(120)   a. Nag-atubili siya-ng   h[um]iram   ng pera   sa banko
             AP-hesitate he(PIV)-LNK [AP]-borrow OBJ money DIR bank
             'He hesitated to borrow money from a/the bank'

         b. Nag-atubili siya-ng   hiram-in   ang pera   sa banko
             AP-hesitate he(PIV)-LNK borrow-OP PIV money DIR bank
             'He hesitated to borrow the money from the bank'

In (120a), the actor of *hiram* 'borrow' is pivot, as revealed by the AP morphology on the verb, and the absence of an overt ANG-phrase in the complement. In (120b), the object is pivot, but the actor is still ellipsed. Therefore both pivot and non-pivot actors can be ellipsed.

An object or other non-actor does not normally undergo ellipsis, even if it is the pivot:

(121)   a. Gusto ni   Juan   sun-in   siya   ng doktor
             want ACT John(LINK) examine-OP he(PIV) ACT doctor
             'John wants the doctor to examine him'

         b. *Gusto ni   Juan   sun-in   ng doktor
              want ACT John(LINK) examine-OP   ACT doctor

<div align="right">Schachter (1977:295)</div>

Furthermore, ellipsis of actors is not restricted to true agents: non-agentive actors of various sorts may be ellipsed, even if they are not pivots:

(122) a. Masagwa   ang t[um]a-tanda
           disagreeable PIV [AP]-IMPERF-become-old
           'It is disagreeable to become old'

       b. Gusto niya-ng   g[um]anda
           want   he/she(ACT)-LNK [AP]-beautiful
           'She wants to become beautiful'

       c. Gusto ko-ng   t[um]anggap ng   gantimpala
           want   I(ACT)-LNK   [AP]-receive OBJ prize
           'I want to be the recipient of the prize'

       d. Gusto ko-ng   ma-tanggap ang gantimpala
           want   I(ACT)-LNK   OP-receive   PIV prize
           'I want to receive the prize'

(123) Ayaw   ko-ng   ma-matay sa   Maynila
       not.want I(ACT)-LNK   AP-die   DIR Manila
       'I don't want to die in Manila'

On this evidence, the ellipsis process seems to target actors regardless of whether they are pivots or agents, providing an argument that actors bear a grammatical relation distinct from the pivot. Since this grammatical relation expresses A and s functions, it is a subject.

Kroeger (1993) finds some issues with these generalizations, but nonetheless confirms that susceptibility to complement subject ellipsis is a genuine property of actors, regardless of whether or not they are pivots. And he develops another extremely important point by showing that the non-pivot actor and non-pivot object are core rather than oblique arguments.

Kroeger (1993:40–8) presents three main arguments to the effect that in the non-AP forms, the actor does not become an adjunct or oblique argument, but remains a core argument. Here I will present one, the Participial Adjunct construction, which also shows that the object in the AP constructions is a core argument, and thus provides evidence against the antipassive analysis of actor-pivot constructions as well as the passive analysis of object-pivot constructions, and therefore shows that neither construction should be considered as 'derived' from the other.

Participial adjuncts are clauses introduced by the particle *nang*, which express action simultaneous with that of the main clause, with the subject suppressed, but understood as coreferential to an actor or object argument of the main clause, regardless of focus. Coreference with a dative or prepositionally marked argument is not allowed:

(124)

a. B[in]isita   ni   Juan ang hari nang nag-iisa
   [PERF]-visit(OP) ACT Juan PIV king ADV AP.IMPERF-one
   'Juan visited the king alone [either Juan or the king is alone]'

b. B[um]ista   si   Juan sa   hari nang nag-iisa
   [AP.PERF]-visit PIV Juan DAT king ADV AP.IMPERF-one
   'Juan visited the king alone [only Juan is alone]'

c. H[in]uli   ng   polis   ang mgananakaw nang pumapasok
   PERF-catch(OP) ACT police PIV thief   ADV AP.IMPERF:enter
   sa   banko
   DAT bank
   'The police caught a/the thief entering the bank [either thief or police
   are entering]'

d. Nang-huli   ng   mgananakaw nang polis   nang pumapasok
   AP.PERF-catch OBJ thief   PIV   police ADV AV.IMPERF:enter
   sa   banko
   DAT bank
   'The police caught the thief entering the bank [either thief or police
   are entering]'

The non-ambiguity of (b) shows the difference beween the (non-core) dative, which can't be understood as the subject of the *nang* construction, and the arguments marked by *ng* and *ang* in the other examples, which can be. This, together with Kroeger's other tests, establishes a core–oblique divide with *ng* phrases on the core side, regardless of whether they are objects of AP verbs or the agents of OP ones.

   So we have a situation where actor and object are core arguments regardless of whether they are pivot or not, and where furthermore the actor is a privileged target for complement subject ellipsis, and also outranks other arguments on a hierarchy relevant for reflexivization (an actor can reflexivize anything, and nothing can reflexivize it). The concept of core vs oblique arguments seems supported, but one of the core arguments has subject-like properties regardless of whether or not it is the pivot. In terms of the ideas introduced at the beginning of the section, the actor will be the a-subject, the pivot the p-subject. Therefore in AP sentences the a-subject and the p-subject are the same NP, which is also the A/S (giving a sentence structure similar to that found in nominative–accusative languages), but in OP sentences the P is p-subject while the A is the a-subject, giving a grammatical structure similar to what is found in syntactically ergative languages. This analysis thus provides the useful properties of the passive and antipassive analyses without suffering from their drawbacks.

Splitting the subject grammatical relation into a-subject and p-subject there-
fore helps to elucidate the Philippine type as well as syntactic ergativity, both
mixed and unmixed. The difference between these types of languages and more
familiar languages such as English is that in the latter there is only one subject-
like grammatical relation rather than two, with the sole subject-like relation
tending to have the typical properties of both a-subject and p-subject (one
could think of both kinds as existing, but always being the same NP).

4.3   *The universal status of a- and p-subjects*

We have now seen that some languages have a 'full' subject combining the
properties of a- and p-subjects, whereas others split them into two distinct
grammatical relations. A further question is whether these two kinds of subject
are always found, whether individually or combined. The answer appears to be
that p-subjects are clearly not universal, while the issue is rather doubtful for
a-subjects.

The languages without p-subjects would be languages such as Warlpiri, which
lack passive or antipassive rules that alter the semantic role of the NP in a recog-
nizable pivot position (subject for nominative–accusative languages, absolutive
for languages with ergative syntax). Although formally inclined linguists have
tended to neglect the different significance of a putative subject relation in lan-
guages with and without a passive rule, it has been discussed at some length
in Van Valin (1981), Foley and Van Valin (1984), and Van Valin and LaPolla
(1997:265–6); see also Foley chapter 8). It is hard to avoid the conclusion that
if a language lacks any rules altering the semantic role of an NP in 'subject'
position, the significance of that position in the functioning of the language
must be different from that of a similar position in a language that has such
rules.

Presence vs absence of a p-subject provides a straightforward account of the
difference, and has been what has been proposed since Foley and Van Valin
(1984), under various terminologies. P-subjects, for example, are frequently
preferred, or required, to be definite, but, clearly, no such requirement can
plausibly exist in a language without passives, where traditionally recognized
subjects would be a-subjects.

What about absence of a-subject? One possible case is Dyirbal, where there
is no clear and compelling evidence for grouping s and A together, but only s
and P. So Dyirbal might be a language with a p-subject, but no a-subject, and
the same would be true of other 'pure syntactic ergative' languages, if these
exist. But, after more than twenty-five years, Dyirbal is still the only reported
case of a pure ergative language that has withstood scrutiny. Furthermore the
data on Dyirbal is limited, and there is little prospect of getting additional data
relevant to the question of whether or not it has a-subjects.

Another potential source of languages without a-subject is languages that have been argued to lack grammatical relations entirely. This claim has been made for a number of languages, including Manipuri and Kannada (Bhat (1991)), and Chinese, Archi, and Acehnese (Van Valin and LaPolla (1997)), on the basis that grammatical phenomena in these languages are controlled directly by semantic roles and pragmatic functions, rather than requiring an intermediate system of grammatical relations. While these claims are very interesting and worthy of being taken seriously, I don't think they are fully established yet. We will here consider Manipuri, and then the phenomenon of 'split intransitivity', which raises similar questions about the role of grammatical relations.

4.3.1   Manipuri   Manipuri, spoken in India, Myanmar and Bangladesh has NP-markers which Bhat labels as 'nominative' and 'accusative' case, although their use departs somewhat from what is usual for cases with these names. Nominative can be found on transitive and intransitive putative subjects, and accusative on putative objects:

(125)   a. Ma-nə   əy-bu   kawwi
             he-NOM   me-ACC   kicked
             'He kicked me'

         b. Ma-nə   kəppi
             he-NOM   cried
             'He cried'

However the nominative is omitted from presumed A/S when these are not expressing volitionally controlling participants:

(126)   a. Ma   əy-bu   uy
             He   me-ACC   saw
             'He saw me'

         b. Ma   sawwi
             He   angry
             'He is angry'

Some verbs appear to require or forbid the use of *nə*, while for others usage varies depending on whether the verb is expressing intentional activity or not (Bhat (1991:119–20)). The suffix has some additional uses which are interesting, but not relevant to this discussion.

Bhat (1991:123) describes the use of the accusative marker as follows:

(127)   a. the referent of the marked noun phrase must be animate
         b. some effect must have been produced on it by an external agency
         c. it must be involved in an action or a process (and not a state)

Example (128a) below is a clear case of an affected argument meeting this description:

(128)   a. ma-nə   huy-bu   kawwi
          he-NOM   dog-ACC   kicked
          'He kicked the dog'

       b. ma-nə   tebəl   kawwi
          he-NOM   table   kicked
          'He kicked the table'

The (b) example lacks the marker because the affected object is inanimate. However, it is unexplained why 'see' (126a) classes its 'seen' argument grammatically as if it was affected.

   With some verbs, the presence of the accusative marker *bu* seems to indicate a more active as opposed to less active version of the event:

(129)   a. əy ma-bu   sawwi
          I   him-ACC   angry
          'I am angry with him [showing anger]'

       b. əy ma   sawwi
          I   him   angry
          'I am angry with him'

But it is unexplained why 'see' classes with the overt display of anger rather than the mere existence of emotional state.

   In addition to marking what might be regarded as somewhat generalized patients, *bu* can under certain circumstances mark recipients of verbs of giving, and what are sometimes called 'causee agents' of causative verbs, that is, participants who are acted upon by the instigating agent of a causative and then produce the effect described:

(130)   a. əy-nə   ma-bu   sel   pi
          he-NOM   I-ACC   money   gave
          'He gave money to me'

       b. əy-nə   ma-bu   tebəl   ilhəlli
          he-NOM   I-ACC   table   caused.to.push
          'He made me push the table'

These uses are subject to the restriction that none of the other arguments of the verbs be able to take *bu*; if this condition isn't met, the locative *də* is used instead, and can be used in any event (so both instances of *bu* in (130) could be replaced with *də*).

There are several complexities in this system which we can't look at here:

a. The predictable uses of *nə* and *bu* are optional. The full circumstances are not entirely clear to me from Bhat's discussion, but it seems that, for example, any of the case-markers in (128) could be omitted.

b. The markers have additional uses to indicate strictly pragmatic functions, in which case they are placed after instances of the markers that are signalling semantic roles (Bhat (1991:126–30)).

Now, considering the issue of a-subjects, a proponent of the universality of grammatical relations could suggest that the marker *nə*, in its function as a semantic role marker, applies only to A/S, that is, a-subjects (since this language has no passive rules), therefore providing some evidence for the relevance of an a-subject concept. The counter-argument is that the distribution of the *nə* marker can be characterized in purely semantic terms, along the lines of 'instigating and intending agent' (someone who does something because they want to do it). This would be expected to prevent *nə* from appearing on a causee agent, because such an agent is being described not as doing something because they want to, but because the instigatory agent makes them do it. So it is certainly plausible that the distribution of *nə*, insofar as this is related to semantic roles, might be determined directly by its semantic role of volitional agent, rather than involving a-subject or other grammatical relations as an abstract intermediary.

Similarly, the distribution of *bu* might well be determined by a semantic role, although it is not so clear from the evidence given exactly what that role would be. But an indication that a semantic role rather than a grammatical relation might be the crucial factor is provided by certain negative sentences, which can have *bu* rather than *nə* on their agents:

(131)   a. ma-bu   laktre
           he-ACC   came.not
           'He didn't come'

        b. layriksi   əy-bu   padri
           book.this he-ACC   read.not
           'He didn't read this book'

Bhat suggests that the accusative is motivated by an implication that some outside influence affected the agent, preventing them from performing the action (Bhat (1991:122–3)). This is evidence that the distribution of *bu* is determined by a semantic role along the lines of 'something that is influenced', rather than by a grammatical relation such as 'object'.

Bhat considers various other phenomena beyond case marking which might involve grammatical relations in Manipuri, and finds no evidence

that they do. For example, there is a participial construction which doesn't allow both clauses to contain non-coreferential actors/causers, but does allow both clauses to contain coreferential actors/causers, only one of which is expressed:

(132)   a. Ra:ju akki tandu   be:yisida
          Raju rice brought(PAST.PTCPL) cooked
          'Raju brought the rice and cooked it'

        b. *Ra:ju akki tandu   hari be:yisida
          Raju   rice brought(PAST.PTCPL) Hari cooked
          'Raju brought the rice and Hari cooked it'   Bhat (1991:75)

However, if only one clause contains an actor/causer, or neither clause does, then no coreferential argument is required:

(133)   a. avanu   be:gane bandu   namage tondard a:yitu
          he(NOM)   early   came(PAST.PTCPL) us(DAT) trouble became
          'We were troubled by his coming early'

        b. mara biddu   ma:du muiyitu
          tree fell(PAST.PTCPL)   roof broke
          'The tree fell and the roof broke'

Obligatory ellipsis and understood coreference of an argument frequently provides evidence for a grammatical relation, but not in this case, because the constraint against this construction of having non-coreferential actors/causers appears to be statable in entirely semantic terms.

   The conclusion is that no grammatical relations at all, including a-subject, are required to describe the grammatical structure of this language. Although this is a very interesting result, it is important to keep in mind that it is inherently difficult to prove a negative, and a few dozen pages of a single investigator's work can't provide conclusive proof that grammatical relations truly play no role at all in the language.

   An example of a potential issue might be whether you could say something such as:

(134)   ma-bu ləppi
        he-ACC cried
        'He cried [because of something somebody did to him]'

If this is acceptable, then the account of the accusative cases in (131) would be corroborated. If not, then it might be a problem to devise a meaning for -*bu* that allowed (131) while excluding (134), and, consequently, there might be a role

for a-subject in Manipuri, for example in the form of a constraint to the effect that a-subjects in positive sentences can't be marked with *bu*.

It is thus not fully established that Manipuri truly lacks grammatical relations, but it is clear that further detailed investigation of the semantic concomitants of the case-marking and other grammatical phenomena ought to lead eventually to a definite answer (and such investigation is possible, since the language is not endangered).

4.3.2   Split intransitivity   One of the many interesting features of Manipuri is the capacity of the accusative marker to appear on a putative subject, as in (131). There turns out to be a considerable number of languages where some intransitive verbs take sole arguments which resemble A in their marking or grammatical behaviour, while others take sole arguments resembling P. This phenomenon, called 'split intransitivity', 'split-s marking', or 'unaccusativity', is widespread in the Americas, also occurring in languages of Indonesia, such as Acehnese (Durie (1985)), and, it turns out, in a somewhat subtle form, in many European languages. For excellent discussion of split intransitivity see Foley, chapter 7, section 1.4, and Dryer, chapter 4, section 2.4.2. Split intransitivity is easy to recognize, although the best way analyse the languages exhibiting it is not always clear.

A fairly typical example is Choctaw (W. Davies (1986:14–16)), originally a language of Mississippi). In this language, A and P are cross-referenced with distinct series of affixes (some prefixes, others suffixes):

(135)   a. Chi-bashli-li-tok
             2(ACC)-cut-1(NOM)-PAST
             'I cut you'
         b. Is-sa-sso-tok
             2(NOM)-1(ACC)-hit-PAST
             'You hit me'

One of the two main types of intransitives takes the 'nominative' (A) agreement:

(136)   a. Hilha-li-tok
             dance-1(NOM)-PAST
             'I danced'
         b. Ish-īpa-h-õ
             2(NOM)-eat-PRED-Q
             'Have you eaten?'

These are verbs whose sole argument (s-function) NPs have agent-like semantic roles.

The other main type takes the 'accusative' (P) agreement:

(137)   a. Sa-hohchafo-h
            1(ACC)-hungry-PRED
            'I am hungry'
         b. Chi-cha:ha-h
            2(ACC)-tall-PRED
            'You are tall'

These are verbs whose S arguments participate in various kinds of involuntary states and events.

The markers used to cross-reference the agent-like S NPs (136) are the ones that are also used for A, while those used for the non-agent-like S NPs of (137) are the same as those used for P, as can be seen by looking at (135). The existence of these two types of intransitive verb is an instance of split intransitivity (there is also a third, small, class of intransitive verbs carrying the markers normally used for recipients, but we will not consider them here). On the basis of this it is reasonable to call the first kind of S 'S$_A$' (S with significant resemblances to A), the second 'S$_P$' (S with significant resemblances to P).

What is really behind this and other instances of split intransitivity is, however, not so clear. A conclusion that one might start to draw from the data so far is that this language has direct reference to semantic roles, reflecting some kind of agent/patient distinction, and that grammatical relations are consequently unnecessary, as Bhat argues for Manipuri. But unlike Manipuri, there is a further coding feature whereby A is treated the same as all S regardless of their semantic role or choice of cross-reference marker. This is nominal case marking.

If an A/S argument is expressed as a full NP in choctaw, then it appears in the nominative case, marked by the ending -*at*, expressed as -*at/-yat/-t*, whereas full NPs with other grammatical functions optionally take the oblique marker -*yã*, regardless of their cross-referencing on the verb:

(138)   a. Ofi-yat   towa(-yã) lhioli-tok
            dog-NOM   ball(-OBL) chase-PAST
            'The dog chased the ball'

         b. Issoba-yat   ĩpa-tok
            horse-NOM   eat-PAST
            'The horse ate'

         c. Chim-alla-t   cha:ha-h
            your-child-NOM tall-PRED
            'Your child is tall'

The case marking on NPs thus reflects a unitary s category, and treats it the same as A, in spite of the split treatment of agreement. Choctaw doesn't have a passive, which shows that the basis for identification of A and s is not that they are both p-subjects. We conclude that Choctaw has an a-subject category, in spite of the split in intransitive predicates, since the two kinds of s show behaviour in common (NP marking) as well as differences (cross-referencing), and the common behaviour furthermore cannot be attributed to p-subject because the language has no passives, and therefore lacks p-subjects. Choctaw therefore conforms to the generalization noted by Dixon (1994:75) that there is almost (but not quite) always evidence that the two kinds of intransitive subjects should be grouped together as some kind of single grammatical relation, in spite of their differences (one of the exceptions is Acehnese, to be discussed below).

Split intransitivity has long been known as a feature of 'exotic' languages, but one of the more significant linguistic discoveries of the late '70s and early '80s is that it is also quite common, in a somewhat subtle form, in European languages, where it is generally known as 'unaccusativity'. In unaccusativity, s_A and s_P are superficially the same in terms of coding features, but more careful consideration of syntactic properties reveals differences, with s_A resembling A, and s_P resembling P.

This was demonstrated extensively for Italian by Perlmutter (1983). In this language there are two kinds of intransitive verbs, some taking *avere* 'have' to form a past tense, the others taking *essere* 'be'. In either case the NP in s function can appear before or after the auxiliary and the main verb:

(139)   a. Due persone sono rimaste
            two   people   are   remained
            'Two people remained'

         b. Sono rimaste   due persone
            are   remained   two people
            'Two people remained'

         c. Due persone hanno reagito
            two   people   have   reacted
            'Two people reacted'

         d. Hanno reagito due persone
            have   reacted two people
            'Two people reacted'

The semantic basis of this split has been a matter of debate; early authors such as Perlmutter (1983) argued that there wasn't any consistent one, while Van Valin (1990) argued that it was aspectually based: 'telic' verbs with a definite result state taking *sono*, those without taking *avere*.

Although all of these s are superficially similar (for example the finite verb agrees with them), there are a variety of syntactic differences. For example, for verbs taking *essere* 'be' as their auxiliary, when the s is postverbal, there can be a partitive clitic before the verb, applying semantically to a quantifier in the postverbal s position. This is not possible for verbs taking *avere* 'have':

(140)   a. Ne   sono rimaste   due
           of.them   are remained   two
           'Two of them remained'

        b. *Ne   hanno reagito due
           of.them   have reacted two
           'Two of them reacted'

One might imagine that there is just a constraint that *ne* 'of them' cannot be used with the auxiliary *avere* 'have', but in fact it can be, to apply to the P of a transitive verb:

(141)   Giorgio ne   ha   comprate due
        George of.them   has bought   two
        'George bought two of them'

What appears to be happening is that *ne*-cliticization is a property of P that is shared by postverbal s_P but not by s_A (also of course not by A). This is one of a number of phenomena whereby A and s_A seem to be similar, and opposed to P and s_P. In spite of its greater subtlety, 'unaccusativity' in European languages seems to be the same phenomenon as the more obvious and longer-known cases of split intransitivity, and is recognized as such in Foley, chapter 7, section 1.4.

There have been a variety of theoretical proposals about the nature of split intransitivity, typically involving arrangements whereby s_P shares some structural relationships with P, and s_A with A. In addition to Perlmutter (1983), see Marantz (1984), Burzio (1986), L. Levin (1988), Zaenen (1993) and Van Valin and LaPolla (1997) for a representative sample. However, there is another possibility, which is that the distinction involves direct sensitivity to semantic roles.

In early work this possibility was discounted, due to difficulties in identifying exactly what semantic role was involved, but more recent investigations, such as Van Valin (1990), Mithun (1991), and B. Levin and Hovav (1995), have tended to find an increasing degree of semantic regularity. A small number of related semantic distinctions seem to be involved, such as whether the verb involves activity (as opposed to describing a state); whether the action is volitional; or whether it is 'telic', having a definite endpoint, as opposed to indefinitely continuous.

A semantic basis for the split has been specifically argued for Acehnese, a language of Sumatra in Indonesia, by Van Valin and LaPolla (1997:255–60), on the basis of work by Durie (1985, 1987, 1988). In this language, A/s_A take an obligatory proclitic, illustrated in (142a,b), while P/s_P take an optional enclitic, illustrated in (142a,c):

(142) a. gopnyan ka   lôn-ngieng(-geuh)
           (s)he(P) already 1-see(-3)
           'I saw him/her'

       b. gopnyan geu-jak
           (s)he   3-go
           '(S)he goes'

       c. gopnyan rhët(-geuh)
           (s)he   fall(-3)
           '(S)he falls'

However, unlike the case in Choctaw, there is no clear evidence that s_P has significant properties in common with s_A and A, and hence no clear case for the existence of an a-subject grammatical relation.[23] There are, however, various grammatical phenomena applying to A/s_A, and others to P/s_P, but none to s_A/s_P/A.

For example the verb *tém* 'want' requires its complement to have an A-like subject, which furthermore cannot be expressed as an overt NP, nor as a proclitic. A P-like argument is not acceptable, whether it belongs to a transitive or intransitive verb:

(143) a. gopnyan geu-tém (*geu-)jak
           (s)he   3-want   (3-)go
           '(S)he wants to go'

       b. geu-tém (*geu-)taguen bu
           3-want   (3-)cook   rice
           '(S)he wants to cook rice'

       c. *gopnyan geu-tém rhët
            (s)he   3-want   fall
            '(S)he wants to fall'

Acehnese might then be an example of language with split s phenomena but no a-subject. A possible analysis would be to say that it has one grammatical relation associated with A function, and another with P function. Either of these would be available for one-place predicates, depending on the meaning. But

---

[23] However, Asyik (1987) makes a partial case for a unified intransitive subject relation, but doesn't discuss all of the implications of the differences between his treatment and Durie's. It would be very useful for someone to work out and reconcile the differences between the two treatments.

there is also a very strong correlation between a core argument NP being a volitional agent and an A/s$_A$, and not being such an agent and being a P/s$_P$. This raises the alternative possibility that Acehnese does not distinguish core arguments by means of different grammatical functions, but rather that the differences between them are caused by direct sensitivity to semantic roles, as argued by Bhat.

Although Acehnese appears to lack a-subject, and may well lack distinct grammatical functions distinguishing the core argument, it does seem to have a clear distinction between core and non-core arguments, and very likely p-subject as well. In front of the verb there is a special position which Durie calls 'core topic', which can according to Durie be optionally occupied by a single core argument.[24] This is the position occupied by the initial nominal in all of the examples above that begin with an NP, but it can also be left unoccupied, in which case a postverbal agent of a transitive verb is marked with the preposition *lé*:

(144)   lôn-pajoh lé lôn pisang nyan
         I-ate   by me banana   that
         'I ate that banana'

The use of the preposition makes the form look somewhat like a passive, but note that the verb is still cross-referencing the agent, and there are also significant complexities in the use of the marker which we won't discuss here. For a transitive verb, the core topic can be either A or P (it is P in (142a)); if the core topic is p-subject, then Acehnese would be a language in which choice of p-subject is relatively free.

Acehnese is therefore relevant to the two questions of whether a-subjects are universal, and whether split intransitivity involves a structural syntactic distinction or direct sensitivity to semantic roles. It also suggests that posses-sion of a-subject and p-subject might be typologically independent features of languages, with different languages having either, neither, or both. Further investigation of the language will be required in order to get definitive answers to these questions.

We have thus learned a lot about the geographical distribution and seman-tic correlates of split intransitivity, but we still don't fully understand how it articulates with other aspects of grammatical structure, in particular whether it always involves a distinction of grammatical relations, or is at least some-times best explained in terms of direct sensitivity of grammatical phenomena to aspects of meaning.

---

[24] However, Asyik (1987) offers a significantly different treatment.

5   Conclusion

The functions of NPs can be usefully classified into three different types, seman-
tic, pragmatic, and grammatical. Semantic and pragmatic functions can be
expected to exist on the basis of what language does, since they are based
directly on aspects of meaning. Certain kinds of semantic and pragmatic func-
tion, such as agent and topic, turn out to be important for the functioning of many
languages. The status of grammatical functions is different: these are abstract
intermediaries between the meanings and overt forms of sentences. Languages
differ in their organization of grammatical functions, and some languages have
been argued to lack them entirely, instead using more direct ways of signalling
the semantic and pragmatic functions.

Although the typology of the grammatical relations is diverse, there are
recurring principles of organization. One basic distinction is between 'core'
and 'oblique' functions; although this can be subtle and hard to ascertain in
some cases, it appears to almost always be present. A now well-established
parameter of variation is the status of the 'p-subject', a grammatical relation
associated with but not identical to the pragmatic function of topic. Languages
may or may not have a p-subject, and if one is present, it may be preferentially
identified with A or with P, or neither may be preferred (a chart of the resulting
typology appears at the end of section 2.3 of Foley, chapter 7).

Less clear issues are whether languages may lack a-subject, and whether
split intransitivity always has a structural basis, or may be a matter of direct
sensitivity to semantic roles. It may thus be the case that certain languages lack
grammatical relations functioning as abstract intermediaries between meaning
and overt form. These questions are difficult to answer conclusively, because
of the difficulty of proving the absence of something, but nevertheless they are
extremely important: if some languages have grammatical relations and others
lack them, that would mean a profound difference in the mental structures
responsible for language use in different communities, and would therefore
be an extremely important result. Both positive and negative answers to the
question of whether languages have grammatical relations or not must therefore
be considered carefully and critically.

6   Suggestions for further reading

As presented here, the functions of NP comprise semantic roles, pragmatic
roles, and grammatical functions and relations. The most important sources
for semantic roles are Jackendoff (1990) and Dowty (1991); and for pragmatic
functions, Lambrecht (1994). Chapter 7, by Foley, also has much useful dis-
cussion on these topics. For more on the grammatical functions A, S and P see

Dixon (1994), and Comrie (1981) for a very clear application to the Torres Strait Island language Kala Lagaw Ya.

The history of thought on grammatical relations is long and complex. Cole and Sadock (1977) is a classic collection of older papers on this issue, while Marantz (1984) is a good discussion of their status as abstract intermediaries between form and meaning. Dziwirek, Farrell, and Mejías Bikandi (1990) is a large collection of studies investigating grammatical relations in a wide variety of languages from many current theoretical points of view, while M. C. Baker (1988) is an influential presentation of a framework where they are not presumed as primitives, but defined in terms of more basic structural relationships. Bresnan (2001) presents a different framework in which a typologically diverse range of data are analysed under the assumption that grammatical relations are primitives. Manning (1996) formally integrates into a variant of Bresnan's framework the results of much previous work in many frameworks on grammatical relations in ergative languages, and is the most immediate source of the 'a-subject' and 'p-subject' terminology used here. Wechsler and Arka (1998) apply this style of analysis to Balinese, showing how a language that superficially seems to be similar to English or Bantu languages is actually a variant of the Philippine type.

Foley and Van Valin (1984) is a central foundational work for the general approach to grammatical relations pursued here, which is extended to an extremely comprehensive typological study by Van Valin and LaPolla (1997), investigating an enormously diverse range of languages with extensive references to relevant descriptive and theoretical literature.

4   Clause types

Matthew S. Dryer

0   Introduction

There are at least four senses in which one can talk about clause or sentence types in a language. One way is in terms of the distinction between declarative, interrogative, and imperative sentences. This distinction, really one of sentence type, is discussed elsewhere in this volume by König and Siemund in chapter 5. A second sense of clause type is represented by the distinction between main clause and subordinate clause, and among different types of subordinate clauses. Issues related to this sense are discussed in the chapters on subordination, such as vol. II, chapter 2 by Noonan, and vol. II, chapter 4 by Andrews. A third sense of clause type concerns the way the same event or situation can be spoken about, from different perspectives, with grammatical consequences such as voice and pragmatic consequences such as topic and focus. This kind of variation is discussed in chapter 7 by Foley. The fourth sense, the one discussed in this chapter, involves different types of clauses in terms of their internal structure, primarily surrounding different types of predicates. Here, the most basic distinction is between verbal and nonverbal predicates. In much of this chapter, differences in clause type hinge on the part of speech of words serving in predicates, to which chapter 1 by Schachter and Shopen is relevant. Among clauses with verbal predicates, we can make further distinctions based on the argument structure of the verb, including the distinction between transitive and intransitive clauses and finer distinctions. These are discussed in section 2 below. We first examine, in section 1, different types of clauses with nonverbal predicates.

1   Nonverbal predicates

There are three types of clauses with nonverbal predicates whose properties vary considerably across languages. These are adjectival predicates, nominal predicates, and locative predicates. In English, all three of these predicates occur with the copula verb *be*, as in (1).

(1)   a. My dog is *black*
        b. My dog is *a cocker spaniel*
        c. My dog is *in the house*

In all three of these sentences, it is useful to think of the element following the form of the verb *be*, rather than *be* itself, as the real predicate. The verb *be* is more of a function word than a predicate; its function can be thought of as combining with nonverbal predicates to form what is syntactically a verbal predicate. While all three of these types of clauses with nonverbal predicates are similar in English, all employing a form of the verb *be*, it is more common cross-linguistically for languages to treat at least one of these types differently from the other two, and occasionally to treat all three in different ways.

Some languages lack copulas entirely, expressing nonverbal predicates directly. For example, in Muɻinypata (Walsh (1976)), a language isolate spoken in northern Australia, all three types of nonverbal predicates are simply juxtaposed with their subjects, without any verbal element. Each of the three types of predicates illustrated for English in (1) are illustrated for Muɻinypata in (2).

(2)   a. panŋun kanʸi-ka putput
           woman   this-TOP pregnant
           'this woman is pregnant'

        b. paŋu-ka   lawaŋga
           that.REM-TOP wallaby
           'that's a wallaby'

        c. nukunu-ka   ŋaɻa ḍa   wiɻit
           3SG.MASC-TOP LOC   place bed
           'he's on the bed'

Note that it is sometimes important to distinguish clauses with nonverbal predicates from nonverbal clauses. The English sentences in (1) involve nonverbal predicates, but they are not nonverbal clauses, since they contain a verb, the copula verb. Examples like those in (2), however, where no copula is used, are not only clauses with nonverbal predicates, but are also nonverbal clauses.

1.1   Types of copulas

Some comment is necessary about the range of elements that might be termed copulas, forms that are used with nonverbal predicates. Such elements are most commonly verbs, as with the English copula *be*. In some languages, they have grammaticized from verbs with more specific meanings, like 'sit', and still have such meaning in some contexts. For example, in Wambaya (Nordlinger

(1998)), a West Barkly language spoken in northern Australia, the verb meaning 'sit' is also used as a copula with both locative and nominal predicates, as in (3).

(3)   a. mirra ngirr-aji   nganaarra-ni
           sit   1PL.EXCL-HABIT.PAST   Brunette.Downs-LOC
           'we stayed at Brunette Downs'

       b. ini   gi-n   galyurringi mirra
           this 3SG-PROG   water   sit
           'this is water'

In some languages, the words serving the function of copulas are nonverbal. For example, in Nuer, a Nilotic language spoken in Sudan, there are two copulas, one used with singular subjects, the other with plural subjects, that are historically derived from the singular and plural third person pronouns, but which are now used as copulas, even with subjects that are first or second person, as in (4).

(4)   ε   ɣän   dec
       be.SG 1SG   soldier
       'I am a soldier'

The morpheme *ε* functions elsewhere in Nuer as a third person singular pronominal clitic, as in (5), but in (4) it has grammaticized as a nonverbal copula.

(5)   cɔŋ-ε
       dance-3SG
       'he is dancing'

Similarly, in Swahili, a Bantu language spoken in east Africa, the copula is nonverbal. Verbs in Swahili inflect for the person, number, and noun class (NC) of their subject (and in some cases, their object) and for tense, as in (6a), but the copula does not inflect for any of these categories, and takes the invariant form *ni*, as in (6b), showing that it is not a verb.

(6)   a. wa-toto   wa-na-cheza   mpira
           NC$_2$(PL)-child   3PL.NC$_2$-PRES-play ball
           'the children are playing ball'

       b. wa-toto   ha-wa   ni wa-dogo
           NC$_2$(PL)-child   this-NC$_2$(PL)   be NC$_2$(PL)-small
           'these children are small'

In other languages, the element that combines with the nonverbal predicate is phonologically bound as a suffix or clitic to the predicate expression. For example, in Eastern Pomo (McLendon (1975)), a Hokan language spoken in

California in the United States, a clitic is added to adjectives or locative case-marked nouns, when they are used predicatively, as in (7).

(7)   a. báheʔ ̇q̇o·dí-ʔè
           that   good-COPULA
           'that one is good'

       b. ká·y-n̥a-ʔè
           ground-on-COPULA
           'it's on the ground'

Similarly, in Ngalakan (Merlan (1983)), a Gunwinyguan language spoken in northern Australia, a suffix *-me* is added to nouns or adjectives when they are serving as predicates, as in (8).

(8)   ŋaykaʔ goʔye ŋu-mir̥para-me-niñ
       1SG.ABS here   1SG-child-COPULA-PAST.CONTIN
       'I was a child here'

Note that the effect of adding the copulative suffix *-me* to a nonverbal predicate in Ngalakan is to create a word that functions as a verb. Once this suffix is added, the resultant form takes verbal affixes; in (8), the resulting verb takes the first person singular subject prefix *ŋu-* and the past continuous suffix *-niñ*. Copulative affixes are often called verbalizing affixes.

   Muruwari (Oates (1988)), a Pama-Nyungan language spoken in Australia, provides two ways to express adjectival predicates, one with a copula verb, as in (9a), the other analogous to the Ngalakan construction, in which a copulative suffix is added to the adjective, after which it behaves like a verb, taking verbal inflections, as in (9b).

(9)   a. marnta yi-n-ta-yu
           cold   be-REALIS-PAST-1SG
           'I was cold'

       b. marnta-ma-yu
           cold-COPULA-1SG
           'I am cold'

1.2   *Adjectival predicates*

Adjectival predicates in English are nonverbal because English treats adjectives as a distinct word class from verbs. In many languages, however, the words expressing meanings associated with adjectives in English are simply verbs. In such languages, adjectival predicates are thus not a kind of nonverbal predicate, but simply a type of intransitive verbal predicate. For example, in Cree (Wolfart and Carroll (1981)), an Algonquian language spoken in Canada, predicates

expressing adjectival meanings exhibit the same grammatical properties as other verbs. Compare the forms of the Cree word for 'sleep' in (10) with the forms in (11) of the Cree word for 'big', used predicatively. (In (10c) and (11c), there is a prefix for first person and a suffix for first person singular.)

(10) a. nipā-w   b. nipā-wak   c. ni-nipā-n
          sleep-3SG   sleep-3PL   1-sleep-1SG
          'he/she sleeps'   'they sleep'   'I sleep'

(11) a. mišikiti-w   b. mišikiti-wak   c. ni-mišikiti-n
          big-3SG   big-3PL   1-big-1SG
          'he/she is big'   'they are big'   'I am big'

The identical morphology of the forms in (10) and (11) illustrates how the Cree word for 'big' is a verb, like the word for 'sleep'.

   In Lealao Chinantec (Rupp (1989)), an Oto-Manguean language spoken in Mexico, adjectival words take verbal morphology, though they differ from other verbs in some respects, and thus belong to a distinct stative class of verbs. As we would expect since they are verbs, they do not require a copula. Rather surprisingly, however, they *can* take a copula, while still bearing verbal morphology themselves. The examples in (12a) and (12b) illustrate these two possibilities (the raised capital letters represent tones).

(12) a. ʔiᴴhiáʔᴹ gá:ᴹi   naⱽᴴ-maᴹ
          very   big.INAN.3   CLSFR-tree
          'the tree is very big'
      b. ʔiᴴhiáʔᴹ gá:ᴹi   naᴸ-lɨⱽᴴ   naⱽᴴ-maᴹ
          very   big.INAN.3   STAT-be.INAN.3   CLSFR-tree
          'the tree is very big'

   In some languages, some of the words corresponding in meaning to adjectives in English are verbs, while others belong to a separate nonverbal word class of adjectives, and this can affect whether they occur with a copula or not. For example in Slave (Rice (1989)), an Athapaskan language spoken in northern Canada, there is a class of adjectives which require a copula when used predicatively, as in (13).

(13) ʔeyá yá-kı̨lı̨
      sick DISTRIBUTIVE-be
      'they are sick'

However, many other words with adjectival meaning in Slave are simply verbs. For example the word for 'big' is a verb and therefore takes verbal morphology and does not occur with a copula, as illustrated in (14).

(14)   yá-nechá
        DISTRIBUTIVE-big
        'they are big'

In most languages, words with adjectival meaning can be used predicatively either directly or in combination with a copula. A third possibility is provided by Dravidian languages like Malayalam (Asher and Kumari (1997)) and Kannada (Sridhar (1990)), both spoken in southern India, in which adjectives cannot directly be used predicatively, but must first be nominalized, and then are used with a copula, like normal nominal predicates. The examples in (15) illustrate predicative and attributive uses in Malayalam, showing how an adjective bears a nominal suffix indicating gender and number only when it is used predicatively, as in (15a), and not when it is used attributively, as in (15b).

(15)   a. ii   kuṭṭi nalla-van   aaṇə
            this child good-MASC.SG be.PRES
            'this child is good'

        b. nalla kuṭṭi
           good   child
           'the good child'

1.3   Nominal predicates

In English, clauses with adjectival predicates and nominal predicates are similar, both employing the copula verb *be.* There are many languages which are similar to English in this respect, employing a copula verb with both adjectival and nominal predicates. However, there are many languages in which a copula is not necessary with adjectival predicates but is required with nominal predicates. The example in (16a) illustrates how adjectival predicates in Mizo (Chhangte (1989)), a Tibeto-Burman language spoken in northeast India, do not involve a copula verb, but employ a structure analogous to that used with verbal predicates, as in (16b), with a subject pronoun immediately preceding the verb, even when there is an independent noun phrase functioning as subject (the superscript '1' denotes tone).

(16)   a. keel a   thii[1]
           goat 3SG dead
           'a goat is dead'

        b. Dou[1]a a   zuang[1]
           Dova   3SG jump
           'Dova is jumping'

With nominal predicates, however, a copula verb is required, as in (17), with the nominal predicate preceding the subject pronoun and the copula verb.

(17)   ka   aar[1] a[1]   nii
        1SG   hen   3SG   be
        'it is my hen'

There are also many languages in which no copula is used with either adjectival or nominal predicates. The examples in (18) from Gude (Hoskison (1983)), a Chadic language spoken in Nigeria and Cameroon, illustrate clauses with adjectival and nominal predicates.

(18)   a. gusə nə   minə
           short SUBJ woman
           'The woman is short'

        b. nwənwu nə   Kwalii
           chief   SUBJ Kwalii
           'Kwalii is a chief'

In both clauses in (18), the predicate expression occurs at the beginning of the clause, without any marking, followed by the subject. But the clauses in (18) are distinct in form from verbal clauses in Gude, in that verbal clauses normally contain an aspect marker, as in (19).

(19)   agi   adənə nə   Musa ɗəfəna
        CONTIN eat   SUBJ Musa mush
        'Musa is eating mush'

Thus, even in the absence of a copula, clauses with adjectival or nominal predicates may have properties distinguishing them from clauses with verbal predicates.

In most languages in which adjectival predicates occur with a copula, the noun occurs with the same copula, as in English. However, in Purki (Rangan 1979), a Tibeto-Burman language spoken in India, adjectival predicates and nominal predicates occur with different copulas, as in (20).

(20)   a. kʰo rgyalpoik in-min
           3SG king   be-PAST
           'he was a king'

        b. kʰo rɖamo   ḍuk
           3SG beautiful be.PRES
           'she is beautiful'

In (20a), we get a copula verb *in* with a nominal predicate, while in (20b) we get a different copula verb, *ḍuk*, with an adjectival predicate.

Similarly, in Mauka (Ebermann (1986)), a Mande language, there is a copula which is used with predicate adjectives but which cannot be used with nominal or locative predicates, as in (21).

(21) díⁿ   à   tímí
      honey be   sweet
      'honey is sweet'

Nominal predicates occur with the same copula as locative predicates, but they must also occur with a postposition meaning 'like'. Example (22a) illustrates the use of this copula with a locative predicate, (22b) with a nominal predicate:

(22) a. sò   yè   tú   lɔ́
          horse be.at   forest   in
          'horses are in the forest'

      b. sò´   yè   sòò   lé
          horse be.at   animal   like
          'horses are animals'

In Logo (Tucker (1940/1967)), a Nilo-Saharan language spoken in Zaire, there are two copulative suffixes, one that is usually used with nominal predicates and the other with adjectival predicates. Nominal predicates normally occur with a copulative suffix *-e*, as in (23).

(23) ma   ago-e
      1SG man-COPULA
      'I am a man'

Adjectival predicates in Logo normally take a different suffix *-ro*, as in (24).

(24) a'di tovo-ro
      3SG   lazy-COPULA
      'he is lazy'

But this association with nouns and adjectives is not rigid, and it is possible to get each of these suffixes with the other type of complement. The example in (25a) illustrates a nominal predicate with *-ru* (an alternate form of *-ro*) while (25b) illustrates an adjectival predicate with *-e*.

(25) a. mí   kugú-ru
          2SG   thief-COPULA
          'you are a thief'

      b. 'dia alo   tani-e
          only one   good-COPULA
          'one only is good'

The difference between these two copulas is thus apparently a semantic one. One hypothesis consistent with these examples (and other examples cited by Tucker) is that *-e* denotes a more permanent state while *-ro* denotes a more temporary state. Since the properties represented by nouns are permanent more often than those represented by adjectives, we would expect to get *-e* more often with nouns and *-ro* more often with adjectives.

In some languages, nominal predicates exhibit the same grammatical properties as verbal predicates, including relevant morphology. The example in (26) illustrates this for Lango (Noonan (1992)), a Nilotic language spoken in Uganda.

(26)   a. án   à-dáktâl
            1SG   1SG-doctor.HABIT
            'I am a doctor'

        b. à-tíyô
            1SG-work.HABIT
            'I work'

In both sentences in (26), the predicate word bears the first person singular prefix *à* as well as habitual aspect marking, which is realized by high-low tone (marked by ˆ) on the last syllable: in (26a), the form of the noun for 'doctor' in its predicate use is *dàktâl*, while in other contexts it is *dàktàl*. Note that nouns are in other respects quite distinct from verbs in Lango. For example, they occur with distinct nominal plural forms, as in *dàktə̂lê* 'doctors'. Note that when the noun is modified, the same verbal morphology occurs on the noun, as in (27).

(27)   án   à-dáktâl   à   bὲr
        1SG 1SG-doctor.HABIT   REL   good
        'I am a good doctor'

Clauses with nominal predicates referring to the past or future in Lango normally occur with a separate verb, otherwise meaning 'stay', functioning as a copula. For past time reference, the perfective form of this verb is used, as in (28a), while for future time reference, a form of the verb for 'come' is used, followed by the infinitival form of the verb for 'stay', as in (28b).

(28)   a. án   à-bédò   dàktàl
            1SG 1SG-stay.PERF   doctor
            'I was a doctor'

        b. òkélò   bínô   bèdò   rwòt
            Okelo   3SG.come.HABIT   stay.INFIN   king
            'Okelo will be king'

1.4   *Equational clauses versus clauses with true nominal predicates*

There are two types of nominal predicates, though most languages do not appear to treat them distinctly. The two types are illustrated for English in (29).

(29)   a. Nancy is a lawyer
          b. Sally Smith is the head of this department

The predicate in (29a) is nonreferential and can be viewed as denoting the generic kind 'lawyer'. The predicate in (29b) is referential and identifies the individual denoted by the predicate with the individual denoted by the subject. Both types of clauses with nominal predicates are often referred to as 'equational', but strictly speaking, the term is only appropriate to the second of the two types in (29). In true equational clauses, the subject and predicate can be reversed; we can thus reverse (29b) as (30), with the only difference in meaning being a possible difference in topic and focus.

(30)   The head of this department is Sally Smith

Clauses with nonreferential nominal predicates – or true nominal predicates, as I will call them – cannot be easily reversed. If we try reversing the subject and predicate in (29a), with a true nominal predicate, we get the very archaic sentence *??A lawyer is Nancy*. In so far as this is acceptable, *Nancy* is still the subject and *a lawyer* is still the predicate, and its status is the same as *??Tall is Nancy*, where it is clearer that *Nancy* is subject (cf. *??Tall are Nancy and her mother*). True nominal predicates can be thought of as being more like adjectival predicates, denoting properties of the subject: (29a) attributes to Nancy the property of being a lawyer just as *Nancy is tall* attributes to Nancy the property of being tall.

The difference between equational clauses and true nominal predicate clauses usually corresponds in English to whether the nominal predicate is grammatically definite or grammatically indefinite, as in (31).

(31)   a. My dog is the cocker spaniel
          b. My dog is a cocker spaniel

However, the difference between the two predicate noun phrases in (31) is quite different from the difference between definite and indefinite noun phrases in other syntactic contexts, as in (32).

(32)   a. I saw the cocker spaniel
          b. I saw a cocker spaniel

In (32b), *a cocker spaniel* is referential, but refers to a cocker spaniel that is not known to the hearer. In (31b), in contrast, *a cocker spaniel* is nonreferential.

Many languages do not distinguish equational clauses from true nominal predicate clauses. For example, in Kutenai, a language isolate spoken in western Canada and the United States, (33) can have either interpretation.

(33)   ṅin-i   xaxas
        be-INDIC skunk
        'it was a skunk'; 'it was the skunk'

However, Kusaiean (K. D. Lee (1975)), an Austronesian language spoken in Micronesia, distinguishes equational clauses from true predicate nominal clauses by employing a copulative particle *pa* between the two noun phrases in an equational predicate sentence, as in (34a), while not employing any overt marker in a true predicate nominal sentence, as in (34b).

(34)   a. mwet   luti   sac pa   mwet   sacn
           person teach the COPULA person that
           'the teacher is that person'

        b. mwet   sacn muhtwacn se
           person that   woman   one
           'that person is a woman'

The predicate in a true predicate nominal sentence can occur with the indefinite determiner *se* 'one', as in (34b), or without it, as in (35).

(35)   ma   sacn usr soko
        thing that   banana.plant
        'that thing is a banana plant'

A second example of a language distinguishing equational clauses from true nominal predicate clauses is West Greenlandic (Fortescue (1984)), an Eskimo-Aleut language. Equational clauses in West Greenlandic involve a nonverbal copular particle placed between the two noun phrases, as in (36a), while true nominal predicate clauses involve a verbalizing (copulative) suffix on the predicate noun, as in (36b).

(36)   a. Hansi tassa pisurtaq
           Hansi be   leader
           'Hansi is the leader'

        b. illuqarvi-u-vuq
           town-COPULA-3SG.INDIC
           'it is a town'

A third example is Cebuano, an Austronesian language spoken in the Philippines. In fact, Cebuano can be said to represent the difference in meaning in a rather transparent way. Compare the equational clause in (37a) with the true nominal predicate clause in (37b).

(37)   a. ang duktur ang babayi
            TOP doctor TOP woman
            'the woman is the doctor'

        b. duktur ang babayi
            doctor TOP woman
            'the woman is a doctor'

Noun phrases in Cebuano normally require one of a set of noun phrase markers or articles like the so-called 'topic marker' *ang*, which occurs twice in (37a) and once in (37b). The equational clause in (37a) consists of a sequence of two noun phrases *ang duktur* 'the doctor' and *ang babayi* 'the woman', and the order of the two noun phrases can be reversed as in English, again with a difference that can be characterized in terms of topic and focus, except that in Cebuano the first noun phrase will normally be interpreted as the focus, the second one as topic. The equational nature is represented by the fact that both parts are noun phrases. The predicate in the true nominal predicate sentence in (37b), however, consists of just a noun, without any marker like *ang.* Although it can take modifiers, it is not really a noun phrase at all, but rather one of a number of structures that can occur as instances of predicates. Furthermore, any predicate, nominal or verbal, can combine with noun phrase markers like *ang* to form a noun phrase. Given a simple sentence like (38a), for example, we can take a predicate like *nagtawag nakuq* 'was calling me', and combine it with a marker like *ang* to form a noun phrase *ang nagtawag nakuq* with the meaning 'the one who was calling me', which can occur as an argument of the verb, as in (38b).

(38)   a. nag-tawag   ang   babayi nakuq
            SUBJ.FOCUS:DUR-call TOPIC woman   1SG.NONTOPIC
            'the woman was calling me'

        b. babayi ang   nag-tawag   nakuq
            woman   TOPIC SUBJ.FOCUS:DUR-call 1SG.NONTOPIC
            'the one who was calling me was a woman'

The syntactic construction whereby a predicate like *nagtawag nakuq* 'was calling me' can be combined with a marker like *ang* to form a noun phrase is the same construction whereby what corresponds to noun phrases in European languages are formed: a noun phrase like *ang babayi* 'the woman' is formed by combining the marker *ang* with a predicate, in this case the nominal

predicate *babayi* 'woman'. With this as background, it should be clear that *duktur* 'doctor' in the true nominal predicate sentence in (37b) is not a noun phrase, in contrast to *ang duktur* 'the doctor' in the equational sentence in (37a). Cebuano not only clearly distinguishes equational clauses from true nominal predicate clauses, but does so in a way that makes the difference transparent, since, unlike English, where both types of clauses involve noun phrases in predicate position, Cebuano uses what is syntactically a type of predicate expression in true nominal predicate clauses, but a noun phrase in equational clauses.

1.5   *Optional copulas*

Copulas are obligatory with nonverbal predicates in some languages, while in other languages they are not. In some instances, the use of a copula is simply grammatically optional and is not grammatically conditioned. The examples in (39) from Tamang (Mazaudon (1976b)), a Tibeto-Burman language of Nepal, illustrate the word *căca* 'small' functioning as an adjectival predicate with a copula (in 39a) and without a copula (in 39b).

(39)   a. ná-la   tìm   căca   mú-la
           1SG-GEN house small be-INDEF
           'my house is small'

        b. cū   mĕnto căca
           this flower small
           'this flower is small'

   In other instances, the use of a copula is grammatically conditioned. For example, in Sanuma (Borgman (1990)), a Yanomámi language spoken in Venezuela and Brazil, no copula is used in the present tense, as illustrated in (40a), while a copula is used in the past and future tenses, as illustrated in (40b) and (40c).

(40)   a. hisa   sa
           young.man 1SG
           'I am a young man'

        b. palata ti   hösösö ku-o-ma
           rubber CLSFR resin   be-PUNCT-COMPLET
           'it was rubber'

        c. kaikana   te   ku-ki   kite
           headman 3SG be-FOCUS FUT
           'he will be headman'

In Evenki (Nedjalkov (1997)), a Tungus language spoken in Siberia, the copula is obligatory, as in (41a), except in the present third singular, as illustrated in (41b).

(41) a. bi   alagumni bi-che-v
         1SG teacher   be-PAST-1SG
         'I was a teacher'

      b. minngi ami-m   bejumimni (bi-si-n)
         my   father-1SG.POSS hunter   be-PRES-3SG
         'my father is a hunter'

The optionality of copulas can also vary with the type of predicate. For example, in Chalcatongo Mixtec (Macaulay 1996), spoken in southern Mexico, the copula is normally required with nominal predicates, but is optional with adjectival predicates. The example in (42a) illustrates a clause with a copula and a nominal predicate, (42b) illustrates a similar clause with an adjectival predicate, and (42c) illustrates a clause with an adjectival predicate without a copula.

(42) a. ku   ɨ̃   čàà ká̃ʔnũ
         be.POTENT.3 one man big
         'he will be a big man'

      b. ku   sṹkṹ
         be.POTENT.3 tall
         'he will be tall [when he grows up]'

      c. čá̃ʔã xa-lúlí
         dirty NOMIN-small
         'the boy is dirty'

In Kombai (De Vries (1993)), an Awyu language spoken in Irian Jaya in Indonesia, there is a copulative suffix -a that is attached to various kinds of nonverbal predicates. With nominal predicates, it is optional; in (43a), it is used, while in (43b) it is not.

(43) a. mene af-a
         this   house-COPULA
         'this is a house'

      b. mene a
         this   house
         'this is a house'

It is also optional with adjectival predicates, as illustrated in (44), though it is apparently more common not to use it.

(44) a. mofene rubu-khey-a
         that   bad-ADJ-COPULA
         'that is bad'

      b. mofene rubu-khe
         that   bad-ADJ
         'that is bad'

However, it cannot be used if the adjective bears the intensifying suffix *-rabo*, as in (45).

(45) a   mene yafe-rabo
      house this good-very
      'this house is very good'

It is obligatory, however, if the predicate is a personal pronoun expressing possession, as in (46).

(46) mene nuf-a
      this 1SG-COPULA
      'this is mine'

   In Korowai, an Awju language closely related to Kombai, there is a copula verb that cliticizes optionally onto both nominal and adjectival predicates. However, Van Enk and De Vries (1997) report that the copula is usually present with adjectival predicates but 'infrequently present' with nominal predicates. Example (47a) illustrates the typical adjectival predicate, with the copula, while (47b) illustrates the typical nominal predicate, without the copula. (The suffix glossed 'NEAR' in (47a) signifies a time near to the present, either past or future.)

(47) a. nokhu khakhul   khén-telo-felu-ndé
         we   yesterday angry-be-NEAR-1SG.REALIS
         'yesterday we were angry'

      b. yu nggulun-benè
         he teacher-Q
         'is he a teacher?'

1.6   *Locative predicates / existential clauses*

1.6.1   Locative copulas   The third common type of nonverbal predicate is that of a locative expression, as in English *My dog is in the house.* Some languages are like English in employing the same copula with locative predicates that is used with adjectival and/or nominal predicates. The examples in (48) illustrate this for Babungo (Schaub (1985)), a Niger-Congo language spoken in Cameroon.

(48) a. tíi   ŋwāa lùu wúu   ndâa
         father my   be person smithy
         'my father is a blacksmith'

      b. fázɨ kə̂ lùu ŋkèe kə̀jə̂ə
         food this be good very
         'this food is very good'

      c. ŋwə́ lùu táa yìwìŋ
         3SG be in market
         'he is in the market'

The copula verb *lùu* is used with all three kinds of predicates in (48), with a
nominal predicate in (48a), an adjectival predicate in (48b), and with a locative
predicate in (48c).

   It is very common, however, for a different copula to be used with locative
predicates, one that has location as part of its meaning. Such locative copulas
are often best glossed 'be at'. We saw above in (28) that Lango uses a verb
originally meaning 'stay' as a copula with adjectival and nominal predicates
in the past and future tenses. In clauses with locative predicates in Lango, a
distinct locative copula is used, as illustrated in (49).

(49) án dáŋ à-tíê   ì   cùkúl
      1SG also 1SG-be.at:PRES.HABIT   at school
      'I'm also at school'

   Similarly, in Koromfe (Rennison 1997), a Niger-Congo language spoken in
Burkina Faso, nominal and adjectival predicates occur with a copula *la*, as in
(50a) and (50b) (though the order of copula and predicate is different with
nominal predicates from its order with adjectival predicates), while locative
predicates occur with a locative copula wɛ̃, as in (50c).

(50) a. mə la a   jɔ
         1SG be ART chief
         'I am the chief'

      b. də lugni a   bĩnĩ̃ã   la
         3SG cat.PL ART black.PL be
         'his cats are black'

      c. də wɛ̃ dãã̃nɛ
         3SG be.at at.home
         'he is at home'

   Some languages commonly use as locative copulas a set of words that vary
along some more specific spatial dimension. In Diyari (Austin (1981a)), a Pama-
Nyungan language spoken in Australia, for example, locative predicates occur

with one of three verbs, meaning 'sit', 'stand', and 'lie', depending on which orientation fits best, as illustrated in (51).

(51)   a. wiḻa   marapu   ŋama-yi ŋura-ṉi
              woman many.ABS sit-PRES camp-LOC
              'there are many women in the camp'

        b. ŋapa piṇa   pantu-ṉi pada-yi
              water big.ABS lake-LOC lie-PRES
              'there is a lot of water in the lake'

Because of the element of verbal meaning that is in these locative copulas, one could argue that examples like those in (51) do not involve nonverbal predicates, though they represent the way in which Diyari expresses meanings that other languages express by means of nonverbal locative predicates.

   Cebuano commonly employs a number of different words in predicate locative clauses that are related to the demonstratives in the language and that vary, not for orientation, but for proximity to hearer and speaker (as well as for tense), as in (52).

(52)   a. túqa   si   Místir Abáya sa   Amiriká
              there:NOT.NEAR.HEARER:PRES TOP Mr   Abaya LOC America
              'Mr Abaya is in America'

        b. ánhi   siyá   sa   Sibú
              here:NEAR.1PL.INCL:FUT 3SG.TOP LOC Cebu
              'he will be here in Cebu'

Despite the fact that these words vary for tense, they are not verbs. The verbal system in Cebuano lacks a distinction between past and present tense and the distinction between future and nonfuture found with verbs is represented very differently from how it is with these nonverbal locative words.

1.6.2   Existential clauses Clauses with locative predicates as a type of clause overlap with what is a distinct category of clause in many languages, that of existential clauses. Consider the three examples in (53) from Ma'anyan (Gudai (1988)), an Austronesian language spoken in Kalimantan (Borneo) in Indonesia.

(53)   a. inehni naqan hang sungking
              mother be.at   at   kitchen
              'his mother is in the kitchen'

        b. naqan   erang kaulun wawey mawiney hang tumpuk yeruq
              be.at/exist one   CLSFR woman beautiful at   village   the
              'there was a beautiful woman in the village'

c. sadiq   naqan tumpuk eteqen
   olden.time exist   village Eteen
   'once upon a time there was a village called Eteen'

All three examples in (53) involve a verb *naqan* 'be at, exist'. The clause in
(53a) involves a locative expression *hang sungking* 'in the kitchen' and the
verb *naqan* can be considered a locative copula, linking an expression denoting
something to which a location is attributed (henceforth the theme) to a nonverbal
predicate consisting of a locative expression. The example in (53b) is in some
ways similar. Again, it can be considered a locative copula, linking a theme
expression (*erang kaulun wawey mawiney* 'a beautiful woman') to an expression
denoting a location (*hang tumpuk yeruq* 'in the village'). It can also, however,
be viewed as stating the existence of something (a beautiful woman), and can
thus be equally well described as existential. The third example, in (53c), does
not involve a location; here, only the existence of something is stated. We can
say, thus, that (53a) and (53b) are predicate locative clauses and that (53b) and
(53c) are existential clauses. This says that (53b) is both a predicate locative
clause and an existential clause. Ma'anyan is not unusual in using the same
word for a range of functions that includes that of a locative copula and that of
an existential word.

   Characterizing clauses like (53b) and (53c) as existential in that they state
the existence of something is perhaps somewhat misleading. From a discourse
point of view, the primary function of such clauses is apparently to intro-
duce into the discourse a participant that is new to the hearer. The contrast
between (53a) and (53b) thus corresponds to a pragmatic difference of iden-
tifiability, and hence to a grammatical difference in definiteness in English.
The example in (53a) does not state the existence of the mother; this is
presumably presupposed. In that sense, the example in (53a) is not really
existential.

   While both (53a) and (53b) can be characterized as involving a locative
predicate, Ma'anyan is like many languages in using a grammatically dis-
tinct construction when the theme expression is pragmatically nonidentifiable
(indefinite): in (53a), the theme expression precedes the verb, in the normal
position for subjects in Ma'anyan, while, in (53b), the theme expression fol-
lows the verb, in a position in which subjects in Ma'anyan are not normally
found. It is not clear, in fact, whether the theme expression in (53b) should be
considered a subject in Ma'anyan, and a similar question arises for analogous
constructions in many other languages. But whether or not the theme expres-
sion ought to be considered a subject, the construction in (53b) and (53c) can be
characterized as a distinct construction in the language, since the verb occurs at
the beginning of the clause and either the subject follows the verb or the clause
is impersonal (i.e. subjectless).

English is in fact strikingly similar to Ma'anyan in a number of respects. In English, it is possible to have a predicate locative sentence with either a definite or an indefinite subject, as in (54).

(54)   a. The dog is in the garden
        b. A dog is in the garden

However, a more natural way to express the meaning of (54b) is as in (55), with a distinct existential construction.

(55)   There is a dog in the garden

The construction in (55) resembles the Ma'anyan construction in (53b) in that the theme expression follows the verb. A difference is that in English there is a separate word *there* in subject position, and, by most criteria, the word *there* functions as the subject (though the theme expression can still control agreement, as in *There are two dogs in the garden*).

Furthermore, the English construction in (55) is restricted to clauses with indefinite subjects. Analogous clauses with a definite subject, as in (56), are rather different in a number of ways.

(56)   There is the dog in the garden.

While the clauses in (54b) and (55) mean approximately the same thing, the clause in (56) means something quite different from (54a) and is arguably a different construction altogether.

Many languages are like Ma'anyan and English in using two different constructions with locative predicates depending on whether the theme is identifiable or not, with a distinct existential construction being used when the theme is nonidentifiable. For example, in Malayalam, there are two locative copulas, the distinction largely depending on the identifiability of the theme. Contrast the two examples in (57).

(57)   a. kuṭṭi   tooṭṭatt-il   aaṇə
            child   garden-LOC   be.PRES
            'the child is in the garden'

        b. meeʃa   meel pustakam uṇṭə
            table   on   book   exist.PRES
            'there is a book on the table'

The copula *aaṇə* in (57a) is the same copula used in sentences with nominal predicates, as in (58) (and adjectival predicates, which must be nominalized to be used predicatively, as illustrated earlier in this chapter in 15a).

(58)   avan   ṭiiccar   aaṇə
       3SG.MASC   teacher   be.PRES
       'he is a teacher'

The choice between the two copulas in (57) when used with locative predicates depends largely on whether the subject is identifiable or not, as is indicated by the English glosses in (57), so that the copula *uṇṭə* can be characterized as existential. The two constructions also differ in their normal word order: with identifiable subjects, as in (57a), the subject most often comes first, followed by the locative expression, while in the existential construction, the locative more often occurs first, followed by the subject. Asher and Kumari note (1997:99) that the use of the two copulas does not exactly line up with the identifiability of the subject, that it is sometimes possible to use the existential verb with an identifiable subject, as in (59), and suggest some possible factors governing this usage, but they note that the contrast between the two uses does normally hinge on the identifiability of the subject.

(59)   uṇṇi   viiṭṭ-il   uṇṭə
       Unni   house-LOC   exist.PRES
       'Unni is at home'

While there are languages like the ones discussed here which distinguish a predicate locative construction with an identifiable theme from an existential construction with a nonidentifiable theme, there are many other languages in which the same construction is used, whether the theme is identifiable or not. The example in (60) from Mangarayi (Merlan 1982), spoken in northern Australia, can be interpreted either way.

(60)   mawuj   ja-Ø-ṇi   biyaŋgin   ṇa-boŋgan
       food   3-3SG-be   inside   LOC-box
       'there's food in the box'; 'the food is in the box'

Thus Mangarayi can be said to lack a distinct existential construction.

The existential constructions in Ma'anyan and English both use verbs in their existential constructions, though English also uses a distinct existential word 'there', which is more like a pronoun than anything else (as reflected by its use in tag questions: *There's a dog in the garden, isn't there?*), though one with a highly restricted distribution. In some languages, however, the existential construction does not employ a verb, but rather an existential word whose categorial properties make it different from words in other categories. For example in Cebuano, there is an existential word *may*, illustrated in (61).

(61)   may   bir
        exist beer
        'there is [was, will be] beer'

While *may* looks verbal to the extent that it occurs in clause-initial position, the normal position for verbs in Cebuano, it lacks the morphological characteristics of a verb, and since other categories, like nouns, can occur in clause-initial position when they are predicates, the position of *may* does not provide any basis for calling it a verb, and its category is thus somewhat indeterminate.

In some languages, the existential construction lacks an overt existential word and consists of just the NP expressing the theme. This is the case in Tolai (Mosel 1984), an Austronesian language of New Britain in Papua New Guinea, as illustrated in (62), where the noun phrase consists of just an article followed by a compound noun meaning 'famine' (literally 'season of hunger').

(62)   a   kilala-na-mulmulum
        ART season-LINK-hunger
        'there was famine'

1.6.3   Existential clauses for expressing predicate possession   Languages differ considerably in how they express what can be called predicate possession. In some languages, this meaning is expressed with a transitive verb like English *have,* as in (63), in which the possessor occurs as subject and the possessed item occurs as object.

(63)   John has a new car

However, many languages employ predicate locative or existential clauses to express such meanings, with the possessor expressed as some sort of locative. The examples in (64) illustrate this for Igorot Bontoc (Seidenadel (1909)), an Austronesian language spoken in the Philippines.

(64)   a. wodắ nan ốnash   id   Falídfid
          be.at ART sugar.plantation LOC Falidfid
          'there was a sugar cane plantation at Falidfid'
     b. wodắy nan fắkat is   nan ongốnga
          be.at ART nail LOC ART boy
          'the boy has a nail'
          (literally 'a nail is at the boy')

The examples in (64) are almost exactly analogous: both involve the existential word *wodắ(y)* (the difference between the two forms is not significant), both have the order verb-theme-location/possessor; and both use the same locative preposition to mark the location in (64a) and the possessor in (64b) (*is* and *id* are

phonological alternants of the same preposition). In many languages, however, predicate possession clauses resemble existential clauses to some extent but the possessor expression is treated somewhat differently. Compare the two examples in (65) from Kannada (Sridhar (1990)), a Dravidian language spoken in southern India.

(65)

a. ka:ŋgaru:-gaḷu   a:stre:liya:-dalli iruttave
   kangaroo-NOM.PL   Australia-LOC   be.NONPAST.3PL.NEUT
   'Kangaroos live in Australia'

b. nana-ge mu:varu heṇṇu makkaḷu   idda:re
   1SG-DAT three   female children.NOM.PL be.NONPAST.3PL.HUMAN
   'I have three daughters'

The example in (65a) illustrates a predicate locative sentence, with a noun marked with the locative case functioning as a locative predicate. The example in (65b) illustrates a predicate possession sentence, with the possessed element in the nominative case and the possessor in the dative case, similar to (65a), but with dative case rather than locative case. There is also a difference in the preferred order for the two constructions (Sridhar p.c.): while the order with the dative-marked possessor first in (65b) is clearly preferred over one with the possessed element in the nominative coming first, there is less clearly any preference for the order of the two noun phrases in (65a).

A possessor can be added to the existential construction in Cebuano illustrated above in (61), but appears as the grammatical topic (subject on some analyses), as in (66), rather than as some sort of locative expression.

(66)   may bir   si   Lúling
        exist beer TOPIC Loling
        'Loling has/had beer'

Similarly, Ma'anyan can add a possessor to the existential construction illustrated above in (53c), as in (67).

(67)   aku naqan buku
        1SG be.at   book
        'I have a book'

Note that the construction in (67) vaguely resembles the locative predicate construction in Ma'anyan illustrated in (53a) above, in that both have an NP preceding the verb, but the meaning associated with the two verbal frames is completely different, since in (53a) it is the preverbal nominal that is the theme, while in (67) it is the postverbal nominal that is the theme, while the preverbal nominal is the possessor.

A distinct way of using existential constructions for expressing the meaning of 'have' is to express the possessor, not as a locative, but as a possessive modifier of the noun possessed. The example in (68a) illustrates this for Imonda (Seiler (1985)), a language of the Border family of New Guinea, and in (68b) for Kutenai.

(68)   a. ne-na   motorbike kai li-f-me
            2-GEN motorbike Q   lie-PRES-Q
            'do you have a motorbike?'
            (literally: 'does your motorbike exist?')

       b. niʔs pikʼaks   saṅ ʔat ɬin ʔin-s-i   ʔa·kawutɬa-ʔis
           ART earlier.times but HABIT must be-OBV-INDIC teepee-3.POSS
           'but in earlier times they must have had tepees'
           (literally 'but in earlier times, their teepees must have been')

1.6.4   Other types of existential clauses   While we are in general not discussing negative clauses in this chapter, it is worth mentioning negative existential clauses here because, unlike most other types of negative clauses in which some other element is negated, the negation in negative existential words is often an inherent part of the predication itself. In other words, while in the example in (69a) from Quechua (Weber (1989)), there is a negative word and negative suffix modifying a separate existential word (here just the normal copula verb being used existentially), in some languages there is a single negative existential morpheme, as in the Kutenai example in (69b) and in the Malayalam example in (69c); compare (69c) with the affirmative Malayalam examples above in (57).

(69)   a. mana papa   ka-ra-n-chu
            NEG potato be-PAST-3-NEG
            'there were no potatoes'

       b. ɬuʔ-ni   k=¢ikam   ni¢tahaɬ
           not.exist-INDIC SUBORD=come young.man
           'none of the young men came'
           (literally 'The young men who came did not exist')

       c. iviṭe kooleej illa
           here college not.exist
           'there is no college here'

Another class of existential predicates in some languages involve numerals or quantifier expressions denoting quantity with meanings like 'many' or 'few'. In English, numerals and quantifiers do not generally function as predicates. We thus do not generally say *The men in the room were three* but are more likely to express the intended meaning by saying *The men in the room were three in number* or *There were three men in the room.* In some languages, however,

numerals and quantifiers are used freely as predicates. In some such languages, this is because numerals are simply verbs, and hence such clauses are really ones with verbal predicates. This is illustrated in (70) for Kutenai.

(70)   a. ni̵=as-ni   titqat̓
           INDIC=two-INDIC man
           'there were two men'
           (literally: 'the men were two')

        b. taxas yunaqaʔ-ni suyapi
           then   many-INDIC white.person
           'then there were a lot of white people'
           (literally: 'then the white people were many')

The indicative proclitic in (70a) and suffix in both (70a) and (70b) are verbal affixes, illustrating the status of *(ʔ)as* 'two' and *yunaqaʔ* 'many' as verbs. While such clauses are existential, it is misleading to characterize them as nonverbal clauses, since the numeral in these clauses is the predicate. But in some languages, analogous clauses do count as nonverbal clauses in that the numerals are not verbs morphologically, but can still be used as predicates, as in Hanis Coos, an extinct language possibly belonging to the Penutian family that was spoken on the west coast in the United States (Frachtenberg 1922), as in (71).

(71)   a. katᶜEʹmîs hanʟ lᴇ   qaiʟāʹwas
           five   FUT the rollers
           'the rollers will be five (in number)'
           (or 'there will be five rollers')

        b. yûʹxwä û   hī̵ⁱʹme
           two   his children
           'he had two children'
           (literally 'his children were two')

1.7   *Minor types of clauses with nonverbal predicates*

In addition to the three fairly basic types of nonverbal predicates (adjectival, nominal, and locative) covered in the preceding sections, some less common ones do occur in many languages. The examples in this section, as well as the English sentences used as glosses, illustrate various types of the minor nonverbal predicates. The examples in (72) from Wambaya (Nordlinger (1998)) illustrate genitive predicates (not to be confused with what were called predicate possession clauses in section 1.6.3 above, expressing meaning like 'I have money'). (The Roman numeral 'ɪᴠ' in these examples represents a noun class.)

(72) a. yana   ngarrga!
         this.IV.SG.NOM 1SG.POSS.IV
         'this (money) is mine'

      b. bungmanya-nkal   yaniyaga   warnu
         old.woman-DAT.IV that.IV.SG.NOM tobacco(IV).NOM
         'that tobacco is the old woman's'

The form of the predicates in these examples is the same form that would occur if these predicates (*ngarrga* 'mine' and *bungmanyankal* 'the old woman's') were occurring as genitive modifiers of nouns. English is somewhat unusual in fact in having a distinct set of pronouns whose basic use is in genitive predicates (*mine*, *yours*, *his*, *hers*, etc.). In Awa Pit (Curnow (1997)), a Barbacoan language spoken in Colombia and Ecuador, the forms used as pronominal possessive modifiers of nouns can also be used directly predicatively. The use of *ap* '1SG.POSS' as a possessive modifier of a noun is illustrated in (73a), its use as a predicate, followed by a copula verb, in (73b).

(73) a. ap   pimpul
         1SG.POSS leg
         'my leg'

      b. an   yal=na   ap   ka-y
         this house=TOPIC 1SG.POSS be.permanently-NONLOCUT
         'this house is mine'

Many languages do not allow genitive predicates, requiring that the genitive be modifying some nominal element. For example, in Una (Louwerse 1988), a Central New Guinea language spoken in Irian Jaya in Indonesia, one must if necessary repeat the noun denoting what is possessed, as in (74).

(74) a   yina Karba yina
         that food Karba food
         'that food is Karba's (food)'

Other examples illustrating minor types of nonverbal predicates are given in (75) to (78). The example in (75) from Babungo (Schaub (1985)) illustrates a benefactive predicate.

(75) ŋwà'lə̀ ŋwə̀ lùu tɨ́ Làmbí
       letter   this be to Lambi
       'this letter is for Lambi'

The examples in (76) illustrate three sorts of minor nonverbal predicates from Gooniyandi (McGregor (1990)), a Bunaban language spoken in northern Australia: (76a) is a purpose predicate; (74b) is a simulative predicate; and (76c) is a predicate denoting origin.

(76) a. thangarla   moonyjoo-yoo ligidd-woo
          toothbrush tooth-DAT   clean-DAT
          'a toothbrush is for cleaning teeth'

      b. goornboo ngooddoo yoowooloo-jangi
          woman   that   man-like
          'that woman is like a man'

      c. niyaji yoowooloo moolooddja-nhingi
          this   man   Mulurrja-ABL
          'this man is from Mulurrja'

Example (77), from Tamambo (Jauncey (1997)), an Austronesian language spoken on Vanuatu, illustrates a predicate representing what some have called a referential expression.

(77) sora-e   atea niani matai tanume arua
        talk-NOMIN one this about devil   two
        'this story is about two devils'

And (78), from Finnish (Sulkala and Karjalainen (1992)), involves a comitative (or associative) predicate.

(78) hän on   minun   kanssa-ni
        3SG be.3SG 1SG.GEN with-1SG.POSS
        'she/he is with me'

   In some languages, the meaning of 'have' is expressed with a construction like (78): 'A has B' is 'A is with B'. The example in (79) illustrates this in Koyraboro Senni (Heath (1999)), a Songhai language spoken in west Africa.

(79) a   goo-nda   zaŋka hiŋka
        3SG be.at-with child two
        'he has two children'

Note that the copula used in (79) is the one that is used with locative predicates, as in (80a), and distinct from the one used with nominal predicates, illustrated in (80b).

(80) a. a   goo no baa sõhoo da
          3SG be.at there even now EMPH
          'it is still there even now'

      b. a   ga   ti noŋguru suub-ante
          3SG IMPERF be place   chose-PTCPL
          'it was a select place'

2   Verbal predicates

Clauses with nonverbal predicates constitute the exception and are apparently less frequent in usage than clauses with verbal predicates, in all languages. Because further discussion of topics directly related to verbal predicates occurs elsewhere in this anthology, in Andrews (chapter 3), Keenan and Dryer (chapter 6), Foley (chapter 7), and Talmy (vol III, chapter 2), our discussion here of verbal predicates and of types of verbal clauses is in some ways more cursory – relative to the variety found among languages – than our discussion of nonverbal predicates.

2.1   Transitive versus intransitive clauses

The most basic distinction among verbal predicates is perhaps that between intransitive and transitive predicates, the former taking a single argument, the latter two (or more) arguments. In many languages, like English, the distinction can be further described by saying that transitive clauses have objects while intransitive clauses do not. This requires that we distinguish transitive clauses with objects from intransitive clauses with adjuncts, illustrated respectively in (81a) and (81b).

(81)   a. My dog ate *the hamburger*
          b. My dog is sleeping *in the basement*

In English, the distinction between object and adjunct is represented by the fact that adjuncts are usually marked with prepositions while objects are not. In some languages, this distinction is less clearly made grammatically, but is grounded in the idea that objects complete the meaning of the verb in a way that adjuncts do not. Typically, for example, adjuncts can be added in any clause where they are not anomalous. Thus we can add the adjunct *in the basement* to (81a), yielding *My dog ate the hamburger in the basement.*

The grammatical criteria for distinguishing transitive and intransitive clauses may vary considerably from language to language. It is not immediately obvious, for example, whether the verb *qakiʔ* 'say' in Kutenai, illustrated in (82), is transitive or intransitive, whether the complement clause (*k̓um aȼ niʔ pałkiy* 'the woman laughed') should be considered an object or not.

(82)   hu qakiʔ-ni   k'=umaȼ   niʔ pałkiy
          1SG say-INDIC SUBORD=laugh the woman
          'I said that the woman laughed'

But there is a verbal pronominal suffix *-(n)am* in Kutenai which indicates a nonspecific subject, which can be added only to intransitive verbs and not to transitive verbs, as in (83).

(83)   n=uwas-nam-ni
       INDIC=hungry-NONSPEC.SUBJ-INDIC
       'people were hungry'

Crucially, this suffix *can* be used on the verb *qakiʔ* 'say', as in (84), showing
that this verb in Kutenai is intransitive, and thus that the complement clause
with this verb is not an object.

(84)   taxas qaky-am-ni   k=¢   haɫnuxunaqnam-nam
       then say-NONSPEC.SUBJ-INDIC SUBORD=FUT race-NONSPEC.SUBJ
       'then people said that people would have a race'

2.2   *Ergative versus accusative patterns*

The distinction between intransitive and transitive clauses becomes more impor-
tant in languages with ergative case systems, in which transitive subjects and
intransitive subjects occur in different cases, transitive subjects occurring in the
ergative case, intransitive subjects in the same case as objects, the absolutive
case, as illustrated in (85) from Kewa (Franklin (1971)), a Trans-New Guinea
language.

(85)   a. áá   píra-a
           man.ABS sit-PAST.3SG
           'the man sat down'

       b. áá-mé   étaa   ná-a
           man-ERG food.ABS eat-PAST.3SG
           'the man ate the food'

The absolutive case, used for the subject in (85a) and for the object in (85b), is
unmarked, while the ergative case, used for the subject in (85b), is represented
by the suffix *-mé*.

   A cross-linguistically common property of the case system of Kewa is that
the ergative case is overtly marked, while the absolutive case is a zero case. But
there are also languages with ergative case marking in which both ergative and
absolutive are overtly marked. For example, in Roviana (Corston (1996)), an
Austronesian language spoken in the Solomons, there are overt prepositional
markers for both ergative and absolutive, as illustrated in (86).

(86)   a. taloa se   Zima
           leave ABS Zima
           'Zima left'

       b. seke-i-a   e   Zima se   Maepeza
           hit-TRANS-3SG.OBJ   ERG Zima ABS Maepeza
           'Zima hit Maepeza'

A more unusual ergative case-marking pattern is found in Nias (Brown (2001)), an Austronesian language spoken on an island off Sumatra in Indonesia, in that the ergative case is null and the absolutive case is non-null, realized by a mutation at the beginning of the noun. This is illustrated in (87).

(87)   a. mörö n-asu
            sleep ABS-dog
            'the dog is sleeping'

        b. i-'inu   n-idanö   asu
            3SG.ERG-drink ABS-water dog.ERG
            'the dog is drinking the water'

In addition, (in realis mood) verbs bear prefixes representing the ergative argument, the absolutive arguments not being represented on the verb, as can be seen in these examples.

   Because it is not obvious that the notions of subject and object apply to ergative case systems, many linguists compare ergative and accusative systems (ones based on the more familiar subject–object distinction) in terms of three notions A, P, and S, where the S is the single argument of an intransitive verb, the A is the more agent-like argument in a transitive clause, and the P is the more patient-like argument, as indicated for the examples in (88). (The P is often called 'O' instead. By using P we foreground the typical semantic affiliation of this grammatical function with patients in the same way that A shows its typical semantic affiliation with agents.)

(88)   a. *Pat* saw *the cat*
            A   P
        b. *The dog* barked
            S

Note that the A need not be an agent, nor need the P be a patient, as in (88a), in which *Pat* is an experiencer, and not an agent in the narrow sense of something volitionally causing an event, and *the cat* is not really a patient in the narrow sense of something that is affected by the event, but is rather what has been called a stimulus. But languages often treat experiencers in the same way as agents and treat the stimulus of perception verbs in the same way as patients, justifying A and P as categories.

   The difference between accusative languages and ergative languages can be described in terms of how they group A, P, and S. In accusative languages, Ss and As are treated one way (and we call them subjects), while Ps are treated distinctly (and we call them objects). In ergative languages, Ss and Ps are treated the same, as absolutives, and As are treated distinctly, as ergatives. These two possibilities are represented in (89).

(89)
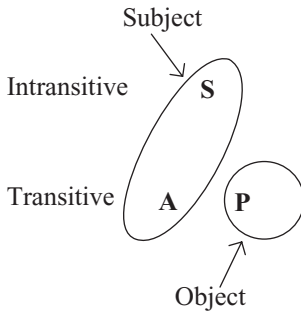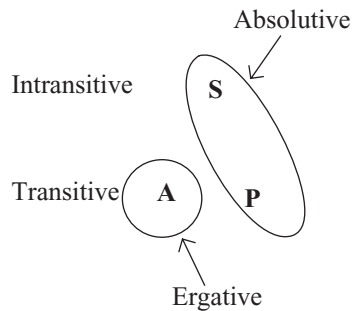
Research has shown that most languages which exhibit ergativity in one part of their system exhibit a more familiar accusative pattern somewhere else in their system. The examples in (85) above show how Kewa exhibits an ergative case-marking system. However, the pronominal affix system on verbs follows an accusative pattern in that the verb inflects for the subject (i.e. S + A). In both examples in (90), the verb inflects for first person singular; in (90a), this is agreement with the S, while in (90b), it is agreement with the A.

(90)   a. ní   píra-wa
           1SG.ABS sit-1SG.PAST
           'I sat down'

        b. né-mé   irikai   tá-wa
           1SG-ERG dog.ABS hit-1SG.PAST
           'I hit the dog'

   See the chapters by Andrews (chapter 3), Foley (chapter 7), and Keenan and Dryer (chapter 6) for further discussion related to ergativity.

2.3   Ditransitive clauses

Some transitive clauses contain two objects, or at least two nonsubject arguments, as in the English sentences in (91).

(91)   a. Nancy gave *Jeff* some flowers
                        R   T
        b. Bob told *Sally*   a story
                     R   T

The noun phrases *Jeff* and *Sally* in (91) are often called indirect objects, the noun phrases *some flowers* and *a story* direct objects. However, because these labels

carry grammatical implications that may not be appropriate for all languages, it is convenient to have more neutral labels for them. By analogy to the notation of A, P, and S, we can use the label 'R' for the recipient-like argument in ditransitive clauses and 'T' for the theme argument (something which undergoes a change in location or to which a location is attributed), as indicated in (91). Semantically, we can say that the R receives the T, either literally, as in (91a) (where Jeff receives the flowers), or metaphorically, as in (91b) (where Sally metaphorically receives the story).

Languages employ a number of different ways of representing the R and the T in ditransitive clauses. English, in fact, has two common constructions, one in which neither the R nor the T is marked with a preposition, and in which the R and T immediately follow the verb, in that order, as in both sentences in (91). In the second construction, illustrated in (92), the T immediately follows the verb and the R occurs later, marked by the preposition *to*.

(92)   a. Nancy gave *some flowers* to *Jeff*
                          T   R
         b. Bob told *a story* to *Sally*
                        T   R

Many other languages employ constructions which are similar to one or the other of these two constructions in English, though it is less common to have both constructions, the way English does. For example, Igbo (Green and Igwe (1963)), a Niger-Congo language of Nigeria, normally uses a construction analogous to the English construction in (91), illustrated in (93).

(93)   o   nyèrè Àdha àkhwa
         3SG gave   Adha   egg
         'he gave Adha some eggs'

In contrast, Ma'anyan (Gudai (1988)) normally uses a construction analogous to the English construction in (92), illustrated here in (94).

(94)   aku ng-amiq   duwit ma ambah-ku
         1SG TRANS-give money to father-1SG
         'I give some money to my father'

These two constructions are particularly common among languages without case affixes and among languages in which the object normally follows the verb.

In languages with case marking on at least one of the two arguments in transitive clauses, we find a number of different patterns of case marking for ditransitive clauses. Two of these patterns are reminiscent of the two patterns found in English. Probably the most common pattern is for the T to be in the

accusative case (which will also be used for the P in monotransitive clauses) and for the R to appear in a separate case, which may be a dative case shared with benefactive noun phrases, or may be some kind of locative case marking used for goal locatives. This is illustrated in (95) for Latin; (95a) shows an intransitive clause, (95b) a monotransitive clause, and (95c) a ditransitive clause.

(95) a. puell-a   vocat
          girl-NOM   call.PRES.3SG
          'the girl is calling'

      b. puell-a   puer-um vīdit
          girl-NOM   boy-ACC see.PERF.3SG
          'the girl saw the boy'

      c. puell-a   libr-um   puer-o   dēdit
          girl-NOM   book-ACC boy-DAT give.PERF.3SG
          'the girl gave the book to the boy'

The notions of 'direct object' and 'indirect object' are useful for characterizing languages like Latin. The category of direct object involves P and T, while indirect objects correspond to Rs.

   Not all languages operate in terms of direct and indirect objects; in other words, not all languages group Ps and Ts together and treat Rs differently. A distinct pattern is found in Kunama (E. D. Thompson (1983)), a Nilo-Saharan language spoken in Ethiopia, in which the R occurs with the same case marking as the P, the T occurring with distinct case marking; the examples in (96) illustrate an object suffix -si marking a P in (96a) and an R in (96b), while the T in (96b) is unmarked.

(96) a. ka   ita-si   intike
          man house-OBJ saw
          'a man saw a house'

      b. dark-oa-m   ikka-si bia   išoke
          woman-that-SUBJ son-OBJ water gave
          'the woman gave water to her son'

   Yoruba, a Niger-Congo language spoken in Nigeria, employs a similar pattern of case marking, though in Yoruba it is the P and the R which are unmarked, while a preposition ni marks the T, as illustrated in (97).

(97) a. Mo ri   baba   e   l'ana
          I   saw father your yesterday
          'I saw your father yesterday'

b. Ajaki ko   Ayo ni   Yoruba
   Ajaki taught Ayo PREP Yoruba
   'Ajaki taught Ayo Yoruba'

c. Mo ya   a   ni   owo
   I   lend him PREP money
   'I lent him some money'

Languages like Kunama and Yoruba can be described in terms of a distinction between primary objects (P + R) and secondary objects (T). Thus, we can say that the object case in Kunama marks primary objects and that primary objects in Yoruba are unmarked, while secondary objects are marked with the preposition *ni*.

The difference between the pattern illustrated by Latin and the pattern illustrated by Kunama and Yoruba can be summarized in the diagrams in (98), similar to those in (89) to distinguish an accusative pattern from an ergative pattern.

(98)   a. direct vs indirect object   b. primary vs secondary object

There is, in fact, a third way in which languages can group the two types of objects, namely treating them all the same way. For example, in Mising (Prasad (1991)), a Tibeto-Burman language spoken in India, there is an accusative case which is used for the P in monotransitive clauses, as in (99a), and for both the R and the T in ditransitive clauses, as in (99b).

(99)   a. bï   kedi-ëm   dɔ́pɔ-duŋ
           3SG mango-ACC eat-PRES
                  P
           'he eats mango'

        b. nɔ-kke   awë-dë   bulu-m   kitab-dë-m   bi-duŋ
           2SG-GEN son-ART 3PL-ACC book-ART-ACC give-PRES
                               R   T
           'your son is giving them a book'

These three possible treatments of ditransitive clauses (the two in (98) plus the possibility illustrated in (99) for Mising) interact with the contrast of accusative and ergative systems to define six possible systems. The three types just discussed illustrate the three possible types in an accusative system (where Su stands for 'subject', Ob stands for 'object', PO stands for 'primary object', and SO stands for 'secondary object').

(100)   Accusative languages:
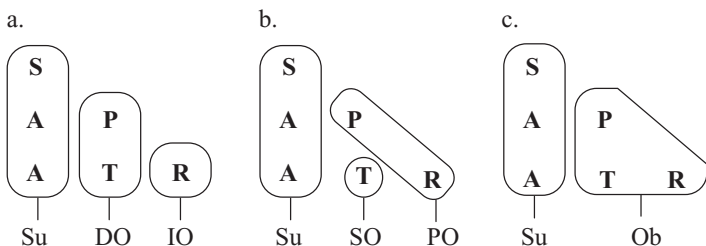
Since the languages used to illustrate the three treatments of objects (Latin, Kunama, Yoruba, and Mising) are all accusative, they also illustrate each of the three patterns in (100), namely (100a) for Latin, (100b) for Kunama and Yoruba, and (100c) for Mising.

There are three analogous possible types of ergative languages: absolutives involve the combination of S and P, but there are three possible ways, analogous to the three types in (100), in which an absolutive category in a language can treat the T and the R. These are given in (101).

(101)   Ergative languages:
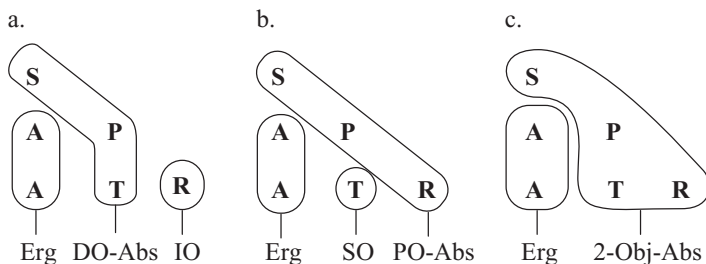
The three patterns shown in (101) involve three different types of absolutives: direct object absolutives, which group Ss with direct objects (P and T), as in (101a); primary object absolutives, which group Ss with primary objects (P and R), as in (101b); and two-object absolutives, which group Ss with monotransitive objects (Ps) and both objects in ditransitive clauses (both T and R), as in (101c).

The case system of Basque (Saltarelli (1988)), a language isolate spoken in northern Spain and southern France, illustrates the direct object absolutive pattern in (101a), as illustrated in (102).

(102) a. katu-a   etza-n-da   d-a-go
          cat-ABS   lie.down-PERF-ADV 3-PRES-be
          'The cat is lying down'

       b. ama-k   gona gorri-a eros-i   d-u-Ø
          mother-ERG skirt red-ABS buy-PERF 3-AUX.PRES-3SG
          'Mother has bought a red skirt'

       c. ni-k   aita-ri   diru-a   eska-tu   d-Ø-io-t
          1SG-ERG father-DAT money-ABS ask-PERF 3-AUX-3SG-1SG
          'I have asked father for money'

The absolutive case is used in Basque for DO-absolutives, i.e. for Ss, as in (102a), for Ps, as in (102b), and for Ts as in (102c); the ergative case is used for As, as in (102b) and (102c); and the dative case is used for Rs, as in (102c).

The case system of Québec Inuktitut (Dorais (1978)), an Eskimo-Aleut language spoken in Canada, is an instance of the primary object absolutive pattern given in (101b) above.

(103) a. Jaani-ø   tikilir-tuq
          Jaani-ABS arrive-PART.3SG
          'Jaani arrives'

       b. Jaani-up   illu-ø   taku-vaa
          Jaani-ERG house-ABS see-INDIC.3SG.3SG
          'Jaani saw the house'

       c. anguti-up Jaani-ø   aitu-paa   illu-mik
          man-ERG   Jaani-ABS give-INDIC.3SG.3SG house-SECONDARY
          'A man gave Jaani a house'

Inuktitut differs from Basque in that the absolutive case is used in ditransitive clauses for the R rather than the T, so that we can say that it is primary objects that occur in the absolutive case in Inuktitut rather than direct objects, the pattern we saw in Basque. And whereas it is the R in Basque that occurs in a distinct case (the dative case), in Inuktitut it is the T that occurs in a distinct case, here glossed 'secondary' (for 'secondary object').

Ngiyambaa (Donaldson (1980)), a Pama-Nyungan language spoken in southeastern Australia, is an example of a language with a two-object absolutive case-marking system of the sort shown in (101c), with an ergative case for As and an absolutive case that is used, not only for Ss and Ps, but for *both* objects in ditransitive clauses. This is illustrated by the examples in (104).

(104)   a. dhibi   bara-nha   balima-ga
            bird.ABS   fly-PRES   sky-LOC
            'birds are flying in the sky'

        b. miri-gu=na   bura:y   gadhiy-i
            dog-ERG=3.ABS   child.ABS   bite-PAST
            'the dog bit the child'

        c. guya=ndu   bura:y   ŋu-nhi
            fish.ABS=2.NOM   child.ABS   give-PAST
            'you gave a child a fish'

The first words in the examples in (104b) and (104c) bear pronominal enclitics that are irrelevant here; for example =*ndu* in (104c) indicates that the subject of the clause is second person. What is crucial here is that the absolutive case in Ngiyambaa follows the two-object absolutive pattern in (101c): the absolutive case is used for Ss, as in (104a), for Ps, as in (104b), and for both Ts and Rs in ditransitive clauses, as in (104c).

2.4   Subtypes of intransitive clauses

The most fundamental division among intransitive clauses is the distinction between intransitive clauses with verbal predicates and clauses with nonverbal predicates, which are generally intransitive. We have dealt with clauses of the latter sort in section 1 above. However, in some languages, there are important further distinctions among intransitive clauses with verbal predicates.

2.4.1   Stative versus nonstative clauses   Perhaps the most common distinction of this sort is a distinction between stative and nonstative verbs, the latter going by various labels such as eventive, process, active or activity verbs. This distinction is a common one in languages in which there is no distinct adjective word class, but in which there is a subclass of verbs whose meaning is typically similar to that of adjectives in languages in which there is a distinct adjective word class.

    For example, in Muna (R. Van Den Berg (1989)), an Austronesian language spoken in Sulawesi in Indonesia, words corresponding to adjectives in other languages are clearly verbal. They take the same inflectional morphology as verbs, as illustrated in (105) and (106).

(105)   a. no-kala   b. no-ghae
            3SG.REALIS-go   3SG.REALIS-cry
            'he goes'   'he cries'

(106)   a. no-ghosa   b. no-kesa
           3SG.REALIS-strong   3SG.REALIS-beautiful
           'he is strong'   'it is beautiful'

The examples in (105) illustrate verbs denoting events with a third person singular realis prefix *no-*, and the examples in (106) show words meaning 'strong' and 'beautiful' inflecting the same way when they occur as predicates.

The example in (107) shows that when a verb modifies a noun in Muna, it must bear participial inflection (PTCPL), consisting of a prefix *mo-* and a suffix *-no*.

(107)   anahi mo-ghae-no
           child PTCPL-cry-PTCPL
           'a child that cries'

The example in (108) shows that when words with adjectival meaning modify nouns they also take the same participial inflection.

(108)   kalambe mo-kesa-no
           girl   PTCPL-beautiful-PTCPL
           'a beautiful child'

These common properties illustrate how words with adjectival meaning in Muna are grammatically verbs.

On the other hand, there are a number of properties that distinguish a subclass of stative verbs in Muna whose meaning corresponds to that of adjectives in languages like English in which a distinct adjective class exists. For example, these stative verbs undergo a morphological process that involves an intensifying prefix *mba-* and reduplication of the verb stem, as in (109).

(109)   no-mba-ghosa-ghosa
           3SG.REALIS-rather-strong-strong
           'he is rather strong'

But the same process is not available for nonstative verbs, as illustrated by the ungrammaticality of (110).

(110)   *no-mba-kala-kala
           3SG.REALIS-rather-go-go

The contrast in (109) and (110) is arguably semantic, since words indicating intensity or degree are often restricted to words denoting states, where the meaning involves a more extreme instance of the state in question, and such a meaning is not directly applicable to words denoting events with meanings like 'go'. On the other hand, there is a second morphological process in Muna that

is restricted to stative verbs, for which a semantic explanation is less clear. This involves a causative prefix *feka-* illustrated in (111).

(111)   no-feka-ghosa-e
         3SG.REALIS-CAUS-strong-it
         'he makes it strong'

Unlike morphemes indicating intensity or degree, causative morphemes are common in other languages with verbs denoting events. However, the prefix *feka-* in Muna does not occur with nonstative verbs, as illustrated by the ungrammaticality of (112).

(112)   *no-feka-kala-e
         3SG.REALIS-CAUS-go-it
         'he makes it go'

There is thus a distinct subclass of stative verbs in Muna and hence a distinct subclass of intransitive clauses.

2.4.2   Split intransitivity   The distinction between nonverbal and verbal intransitive clauses and between stative and nonstative clauses both involve a split among intransitive clauses, but the terms 'split intransitivity' and 'split-S' are commonly applied to splits where intransitive clauses divide into two types depending on whether the single argument (the S) exhibits grammatical properties similar to those of the A in transitive clauses or to those of the P. For example, in Bukiyip (Conrad and Wogiga (1991)), a Torricelli language spoken in Papua New Guinea, there are some intransitive verbs whose S is coded on the verb in the same way as the A in a transitive clause, and other intransitive verbs whose S is coded on the verb like the P in a transitive clause. The examples in (113) illustrate coding on the verb for transitive verbs in Bukiyip.

(113)   a. n-a-la-tú
             NC$_7$.SG-REALIS-build-NC$_{11}$.SG
             'he built it (a house)'

         b. okok kw-a-túl-únú
             she   NC$_4$.SG-REALIS-see-NC$_7$.SG
             'she saw him'

In both examples in (113), the verb bears a prefix indicating the person, number, and noun class of the A and a suffix indicating the same for the P. A singular A belonging to noun class 7 (which is the masculine noun class) is represented by a prefix *n-*, as in (113a), while a singular P belonging to noun class 7 is represented by a suffix *-(ú)nú*, as in (113b). For the majority of intransitive verbs in Bukiyip, the S will be represented on the verb by one of the prefixes

used for As in transitive clauses, as in (114), where we find the same prefix *n-* that occurs in (113a).

(114)   énan n-a-leh
       he   NC7.SG-REALIS-cry
      'he cried'

However, for a minority of intransitive verbs, their single argument is represented by the same set of suffixes that represent Ps in transitive clauses. In (115), for example, we find the single argument represented by the suffix *-(ú)nú* that represented the P in (113b).

(115)   énan élgei-nú
       he   afraid-NC7.SG
      'he is afraid'

The verbs in Bukiyip that behave like *élgei* 'be afraid' are all ones whose single argument is not agentive and whose semantic relation to the verb is in some respects more similar to that of a P in a transitive clause. For this reason, languages like Bukiyip are sometimes described as operating in terms of agent and patient rather than subject and object.

One way to describe split intransitive languages like Bukiyip is in terms of a diagram like that in (116), which contrasts with the accusative and ergative patterns portrayed above in (89).

(116)   split intransitive (split-S) pattern
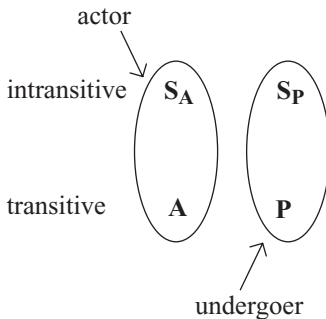
The diagram in (116) splits S into two subtypes, S_A and S_P, where S_A consists of those Ss that behave like As and S_P consists of those Ss that behave like Ps. The term 'actor' is sometimes applied to the union of S_A and A and the term 'undergoer' to the union of S_P and P.

Split intransitive systems most often manifest themselves in the system of pronominal marking on verbs, as in Bukiyip. In some languages, however, it is the system of case marking on noun phrases that operates in a split intransitive

fashion. An example of such a language is Hunzib (H. Van Den Berg (1995)), a Dagestanian language spoken in the Caucasus region in Russia. The example in (117a) illustrates a transitive clause, with overt case marking on the A and zero marking on the P; the example in (117b) illustrates zero marking on an S; the example in (117c) illustrates an S with the same actor marking that occurs on the A in (117a).

(117) a. iyu-l   hɑrɑ b-oho-r
           mother-ACT   cow   NC₄-feed-PRET
           'mother fed the cow'

       b. ože uṫ-ur
           boy sleep-PRET
           'the boy slept'

       c. hɑrɑ-l   heλe-r
           COW-ACT   moo-PRET
           'the cow mooed'

Languages differ as to the basis of the split among Ss. In Hunzib, for example, there are only ten intransitive verbs that take arguments with actor case-marking, while all other intransitive verbs take arguments with undergoer case. These ten verbs are all verbs associated with bodily actions or noises. In most languages, there seems to be some semantic principle related to volitionality or stativity underlying the split, though ultimately the distinction seems to be lexical in that it is not entirely predictable whether the S of a particular intransitive verb will be an S_A or an S_P. While the terms 'actor' and 'undergoer' are sometimes used in a purely semantic sense distinct from the grammatical patterns found in particular languages, the terms are also sometimes used as labels for categories that occur in particular split intransitive languages. Using the terms in this way, we can say that languages differ in the apparent semantic principle underlying the split. In many languages, the split is related to volitionality, volitional arguments appearing grammatically as actors, nonvolitional arguments as undergoers. This is the case, for example, in Choctaw (W. D. Davies 1986), a Muskogean language spoken in the United States. The examples in (118) illustrate the pronominal marking in transitive clauses in Choctaw.

(118) a. chi-pisa-li-tok
           2SG.UNDERGOER-see-1SG.ACT-PAST
           'I saw you'

       b. is-sa-sso-tok
           2SG.ACT-1SG.UNDERGOER-hit-PAST
           'you hit me'

The examples in (118) illustrate two different affixes for first person singular arguments, one a suffix, the other a prefix: the example in (118a) illustrates the first person singular actor suffix *-li*, while the example in (118b) illustrates the first person singular undergoer prefix *sa-*. Intransitive verbs differ as to which of these two affixes occur when their argument is first person singular. The example in (119a) shows a first person singular intransitive actor, represented by the suffix *-li*, while the example in (119b) shows a first person singular intransitive undergoer, represented by the prefix *sa-*.

(119)   a. bali:li-li-tok   b. sa-cha:ha
            run-1SG.ACT-PAST   1SG.UNDERGOER-tall
            'I ran'   'I am tall'

The two verbs in (119) differ in two ways: the verb meaning 'run' in (119a) denotes an event and its argument is volitional, while the verb meaning 'tall' in (119b) denotes a state and its argument is nonvolitional. It is common in languages with a split intransitive system for volitional arguments of event verbs to be actors and for nonvolitional arguments of stative verbs to be undergoers. But languages differ in their treatment of nonvolitional arguments of event verbs. Perhaps the more common pattern is that found in Choctaw, in which such verbs take undergoers, showing that volitionality is the primary semantic factor, as shown in (120), where we find the undergoer prefix *sa-* rather than the actor suffix *-li*.

(120)   a. sa-ttola-tok   b. sa-habishko
            1SG.UNDERGOER-fall-PAST   1SG.UNDERGOER-sneeze
            'I fell'   'I sneezed'

Different semantic factors are apparently relevant to Apurinã (Facundes (2000)), an Arawakan language spoken in Brazil. Actors and undergoers in Apurinã are distinguished by the position of pronominal affixes on the verb: actors are represented by prefixes, while undergoers are represented by suffixes. Both of these are illustrated by the example with a transitive verb in (121).

(121)   n-arika-ru
        1SG.ACT-set.on.fire-3SG.MASC.UNDERGOER
        'I set it on fire'

Intransitive verbs in Apurinã differ as to which of these two sets of affixes they occur with. In (122a), we see an actor prefix representing the volitional argument of an intransitive event verb, while, in (122b), we see an undergoer suffix representing the nonvolitional argument of an intransitive stative verb.

(122)   a. nu-muteka   b. hareka-no
        1SG.ACT-run   good-1SG.UNDERGOER
        'I run'   'I am good'

In Apurinã, the argument of an event verb is invariably an actor, even when nonvolitional, as in the two examples in (123).

(123)   a. nh-iri   b. o-pö-p
        1SG.ACT-fall   3SG.FEM.ACT-die-IMPERF
        'I fell down'   'she died'

Furthermore, there are many stative verbs which take actors rather than undergoers, as in (124).

(124)   a. nu-sãpaka   b. nhi-inhikaka
        1SG.ACT-tired   1SG.ACT-feel.hot
        'I am tired'   'I feel hot'

Further examination would be necessary to determine whether there is some semantic basis to the contrast between those stative verbs in Apurinã which take undergoers and those which take actors.

   Other semantic factors are at play in other languages. In Taba (Bowden (1997)), an Austronesian language spoken in eastern Indonesia, the choice of actor versus undergoer in intransitive clauses is partly sensitive to the animacy of the argument. In Taba, actors are represented on the verb by proclitics, while undergoers are generally not represented on the verb. A simple transitive clause is given in (125), illustrating SVO order for a transitive clause, and illustrating a prefix on the verb for the actor, but no marking on the verb for the undergoer.

(125)   Ahmad n-pun   kolay
        Ahmad 3SG.ACT-kill snake
        'Ahmad killed a snake'

In addition, independent pronouns, if present, precede the verb when they are actors, but follow when they are undergoers, like full noun phrases in these roles. This is illustrated for transitive clauses in (126): (126a) illustrates an independent pronoun *am* 'we, exclusive' functioning as a transitive actor preceding the verb, while (126b) illustrates an independent pronoun *i* 'third person singular' functioning as a transitive undergoer following the verb.

(126)   a. am   a-tala   motor la-we
           1PL.EXCL 1PL.EXCL-meet   boat sea-ESS
           'we (exclusive) met the boat by the sea'

b. ni   mamasi n-wet   i
   3SG.POSS mother 3SG.ACT-hit 3SG
   'his mother hit him'

Analogous intransitive examples are given in (127): in (127a), the argument is an actor and is represented both by an independent pronoun preceding the verb and by a proclitic on the verb, while in (127b), the argument is an undergoer and is represented only by an independent pronoun following the verb.

(127)   a. i   n-tagil   ndara   b. kawail i
            3SG   3SG.ACT-walk too.much   tired 3SG
            'he walks too much'   'he is tired'

In some cases in Taba, the choice of actor versus undergoer is sensitive to the animacy of the argument, as in (128).

(128)   a. n-ha-mlongan   b. ubang da   mlongan
            3SG.ACT-CAUS-long   fence that long
            'he is tall'   'that fence is long'

The verb *mlongan* 'long' in Taba is basically an undergoer intransitive verb, and this use is illustrated in (128b). However, there is a productive process in Taba whereby what is otherwise the causative prefix combines with an undergoer intransitive verb to yield an actor intransitive verb. This is illustrated in (128a), where we find a causative prefix on the verb, but also an actor prefix, illustrating how the argument is an actor in (128a) but an undergoer in (128b). Although both verbs in (128) basically mean 'be long', the one that takes an actor is generally used with humans while the one that takes an undergoer is generally used with nonhumans.

In some languages, issues of split intransitivity extend to clauses with non-verbal predicates. For example, in Kambera (Klamer (1998)), another Austronesian language of eastern Indonesia, nominal and locative predicates are among the class of intransitive predicates that take undergoer clitics for their single argument: (129a) illustrates a transitive clause with a third person actor proclitic *na-* and a third person undergoer enclitic *-ya*; (129b) shows an intransitive clause with the single argument represented by the actor proclitic *na-*; (129c) and (129d) show intransitive clauses with nominal and locative predicates respectively, with their single arguments represented by the undergoer enclitic *-ya*.

(129)   a. na-palu-ya   na   ahu
            3SG.ACT-hit-3SG.UNDERGOER ART dog
            'she hits the dog'

b. na-kapunduh weling la kanjaka
3SG.ACT-jump move.from LOC chair
'he jumps from the chair'

b. hurundandu-ya
soldier-3SG.UNDERGOER
'he is a soldier'

e. la wawa kotak-ya
LOC down village-3SG.UNDERGOER
'he is below the village'

2.4.3 Zero-intransitive (or ambient) clauses Intransitive clauses are usually characterized as involving a single argument while transitive clauses are characterized as involving two (or more) arguments. However, many languages have clauses which can be described as involving *zero* arguments. These clauses are ones that semantically do not involve any arguments, though languages vary as to whether they are treated as lacking arguments in their syntax. This sort of clause normally involves environmental conditions, typically weather conditions. Examples of English clauses of this sort are given in (130).

(130)   a. It is raining
        b. It is cold today
        c. It is hot in this room

In English, clauses like these resemble intransitive clauses like *It is screaming* or *It is weak.* However, in clauses of the latter sort, the subject pronoun *it* is referring to something, most likely nonhuman, while in (130) the pronoun *it* is nonreferential. This is reflected by the fact that the *it* in *It is screaming* can be replaced by some other noun phrase, as in *I am screaming* or *Who is screaming?* But this is not possible in (130a) (**I am raining*, **What is raining?*). The nonreferential *it* in clauses like those in (130) is often referred to as an expletive or dummy subject.

Because of the presence of the expletive subject in the English examples in (130), these clauses are grammatically like intransitive clauses, and though they may be described as being zero-intransitive semantically, they can be described as normal intransitives grammatically. But English is actually rather unusual cross-linguistically in using expletive subjects. This is fairly unusual outside of Europe, although an example of a non-European language that is like English in using a semantically nonreferential third person pronoun in zero-intransitive clauses is Buru (Grimes (1991)), an Austronesian language spoken in Maluku in Indonesia, as illustrated in (131).

(131)   da   deka
          3SG rain
          'it's raining'

Most other languages employ one of four alternative strategies for expressing such meanings.

   Some languages employ a strategy that is similar to English, except that instead of a semantically nonreferential independent pronoun as subject, they employ a semantically nonreferential third person singular form of a verb. As in English, these clauses look like normal intransitive clauses, except that such clauses cannot take an independent noun phrase as subject and the third person singular affix is nonreferential with these verbs but referential with normal intransitive verbs. For example, in Tukang Besi (Donohue (1999)), an Austronesian language of Indonesia, 'it is raining' is expressed by a single word, as shown in (132).

(132)   no-wande
          3.REALIS-rain
          'it is raining'

And just as in English, it is not possible to add any noun as subject, as shown in (133).

(133)   *no-wande   na   wande /langi /lono /'oloo
          3.REALIS-rain NOM rain   /sky   /cloud /day
          'it's raining'

However, unlike English, it is not possible to have a third person singular independent pronoun meaning 'it' as subject, as shown in (134).

(134)   *no-wande   na   ia
          3.REALIS-rain NOM it
          'it is raining'

But the form in (132) does look superficially like a normal intransitive clause expressed entirely by the verb, as in (135).

(135)   no-tinti
          3.REALIS-run
          'she is running'

   A second strategy for expressing the meanings of clauses like those in (130) is by means of a referential subject. In Bukiyip, for example, clauses referring to raining involve a subject noun meaning 'rain', as illustrated in (136).

(136)   echah   h-a-lali
        rain   NC$_{13}$-REALIS-rain
        literally 'rain rains'

In these languages, such clauses are really intransitive (rather than zero-intransitive) since the subject is referential.

A third strategy is employed in languages in which there is no overt expression of any argument, in which there can be no separate pronoun or noun as subject and the verb does not inflect for any argument. An example is given in (137) from Tahitian (Tryon (1970a)), an Austronesian language spoken on the island of Tahiti in the Pacific.

(137)   'e   ūa
        fut rain
        'it will rain'

A fourth strategy is to use a noun meaning 'rain' by itself without any accompanying word, as in Tawala (Ezard (1997)), an Austronesian language spoken in Papua New Guinea, as in (138).

(138)   gadiwewe
        rain(NOUN)
        'There is rain'

This is really an instance of an existential clause of the sort illustrated above in (62) for Tolai, which consists simply of a noun phrase denoting that which exists.

Zero-intransitive clauses in Tolai (Mosel (1984)) superficially look like English clauses in having a third person singular pronoun, as in (139).

(139)   i   ga   bata
        3SG PAST rain
        'it rained'

However, the pronominal word *i* in (139) is an agreement pronoun that is in the verb phrase and not in subject position, in contrast to the *it* in English *it is raining*: this agreement pronoun in Tolai co-occurs with a noun phrase in subject position, as in (140).

(140)   nina ra   tutana i   ga   mait
        that DEF   man   3SG PAST sick
        'that man was sick'

In other words, the agreement pronoun is more analogous to an agreement affix, except that it is a separate word. Thus, this should probably be viewed as a subcase of the strategy illustrated above in (132) for Tukang Besi, in which

there is a pronominal marking on the verb, rather than like English, in which there is a marking in the position of syntactic subjects.

While perhaps the most common sort of zero-intransitive clause are ones involving environmental conditions, many languages also use them for expressions of time, as in the Kutenai and Awa Pit examples in (141a) and (141b) respectively.

(141) a. k=waɫkwayit-s,   ɫa   ɫaxax-i   ɫkam-niṅtik
           SUBORD=be.evening-OBV back arrive-INDIC child-PL
           'when evening came, the children arrived back'

       b. nash-mɨz-i
           be.afternoon-INCEP-NONLOCUT
           'it is getting late'

In both of these cases, a verb is used, but one which does not take any arguments semantically. The verb for 'be evening' in Kutenai in (141a) does take inflection for an obviative subject, a category of inflection that indicates that the subject of that verb is different from the subject of verbs in the surrounding discourse lacking obviative inflection, so that at some grammatical level we may want to say that this clause has a subject; however, this subject can only be realized phonologically by an affix on the verb, not by a separate noun phrase, as is possible with the nonreferential subject *it* in English *it is raining*.

2.5   *Semi-transitive clauses*

The distinction between intransitive and transitive clauses is in principle a straightforward one: intransitive clauses contain a single argument, while transitive clauses contain two or more. In many languages, however, there are some clauses that do not fall easily into one or the other of these two categories, where they behave in some ways like intransitive clauses, but in other ways like transitive clauses. Most often the verbs in such clauses have two arguments semantically, but neither is an agent in the narrow sense of someone or something that volitionally causes the event denoted by the verb.

One sort of semi-transitive clause is found in Koyraboro Senni (Heath (1999)), where objects of transitive verbs normally immediately precede the verb, without any case marking, as in (142). (The second word in (142), *na* 'transitive', occurs in a slot otherwise filled by tense–aspect–mood particles that immediately follow the subject in Koyraboro Senni.)

(142) a   na   ham   ŋaa
       3SG TRANS meat eat
       'he ate meat'

Nonarguments, in contrast, follow the verb and are typically marked with a postposition, as in (143).

(143)　a　koy koyr-aa　　ra
　　　　3SG go　town-DEF.SG LOC
　　　　'he went to the city'

There are a minority of verbs, however, which involve two arguments semantically, but where the nonsubject argument follows the verb, without being marked with any preposition or postposition, as in (144) from Prost (1956).

(144)　a. a　　ga　　ba　　agey
　　　　　　3SG INCOMP like 1SG
　　　　　　'she likes me'

　　　　b. ay　di né　wand-o
　　　　　　1SG see 2SG wife-DEF
　　　　　　'I saw your wife'

　　　　c. ay　ga　　humbur hans-o
　　　　　　1SG INCOMP fear　　dog-DEF
　　　　　　'I am afraid of the dog'

　　　　d. i　ga　　hima　kyer-ey
　　　　　　3PL INCOMP resemble each.other-PL
　　　　　　'they resemble each other'

On the one hand, these clauses are not like normal transitive clauses in that there is no object noun phrase preceding the verb. On the other hand, they are not like typical intransitive clauses: not only do they have two arguments semantically, but the nonsubject argument, though following the verb, is not marked with a preposition. These clauses are neither straightforwardly transitive nor straightforwardly intransitive. A useful label for clauses of this sort is semi-transitive. In a sense they behave more transitively than normal intransitive clauses and less transitively than normal transitive clauses. It is worth noting that none of the subjects in (144) are agents in the narrow sense; the subjects in (144a), (144b) and (144c) are all experiencers of some sort and the subject in (144d) is a patient/theme. The same is true of other verbs in Koyraboro Senni that behave in this way and is a typical feature of semi-transitive clauses in other languages.

　　Djaru (Tsunoda (1981)), a Pama-Nyungan language of northwestern Australia, exhibits a number of distinct types of semi-transitive clauses. The normal pattern for intransitive and transitive clauses is exhibited in (145), with a zero absolutive case for Ss and Ps and an overt ergative case for As.

(145)   a. mawun   jan-an
           man.ABS go-PRES
           'the man goes'

        b. mawun-du ɲaŋ-an   ḍaḍi
           man-ERG   see-PRES kangaroo.ABS
           'a man sees a kangaroo'

However, there is a class of verbs which take two arguments semantically, whose A occurs in the ergative case, but whose P occurs in the dative case, as in (146).

(146)   mawun-du ŋa=la   ḍaḍi-wu   muwu wuŋ-an
        man-ERG   CLITIC.HOST=3SG.DAT kangaroo-DAT search-PRES
        'a man is looking for a kangaroo'

The case marking in (146) is anomalous in that the ergative case is otherwise not possible in Djaru unless there is an absolutive in the same clause.

   A second type of semi-transitive clause in Djaru involves a verb which takes two arguments semantically but which takes one argument grammatically in the absolutive case, the other in the dative or locative cases, as in (147).

(147)   jaɱbagina ŋa=ɲanda   juwa maɲ-an guɲar-a
        child.ABS   CLITIC.HOST=3SG.LOC be.afraid-PRES dog-LOC
        'a child is afraid of a dog'

Clauses like (147) are grammatically identical to intransitive clauses containing a locative nominal in Djaru, differing only in that the locative-marked nominal is semantically an argument of the verb.

   The third type of semi-transitive clause in Djaru is anomalous in that it contains two arguments that are both marked absolutive, as in (148).

(148)   mawun ḍaru   maɲ-an
        man.ABS Djaru.ABS talk-PRES
        'a man talks Djaru'

The absolutive form of *mawun* 'man' in (148) suggests that it is being treated as an S and that the clause is intransitive. The absolutive form of *ḍaru* 'Djaru' suggests that it is being treated as a P and that the clause is transitive. Again the label 'semi-transitive' is a useful label for such clauses.

   Note that with all three types of semi-transitive clauses in Djaru, the meaning of the verb is one that deviates from ones with an agent argument and a patient/theme argument. In the case of the verb meaning 'afraid', one argument is an experiencer and the other argument is a stimulus. In the case of the verbs meaning 'look for' and 'talk', one argument is agentive, but the other argument

is semantically unlike patient/themes in not being affected by the action of the verb.

The third type of semi-transitive clause in Djaru, as in (148), with two absolutive nominals, might be explained in terms of the fact that the absolutive case is a zero case, that *mawun* 'man' is absolutive because it is an S but that *ḍaru* 'Djaru' is zero-marked for some reason other than being grammatically absolutive. But this explanation will not work for certain clauses in Nias. As illustrated above in (87), Nias employs ergative case marking, with an overtly marked absolutive case. But there are a few verbs in Nias that occur with two arguments which both occur in the absolutive case, as in (149), analogous to the Djaru example in (148).

(149)   a. ata'u   n-akhi-gu   n-asu
            be.afraid ABS-younger.sibling-1SG.POSS ABS-dog
            'my younger brother is afraid of the dog'

        b. omasi n-asu   n-akhi-gu
            like   ABS-dog ABS-younger.sibling-1SG.POSS
            'the dog likes my younger brother'

Again, these clauses are quite anomalous in containing two noun phrases in absolutive case, something that is not otherwise possible in the language. They are less transitive than normal transitive clauses in that there is no noun phrase that is case-marked ergative and the verb does not bear a prefix coding either argument, in contrast to normal transitive clauses in Nias, in which the verb bears a prefix for the ergative argument. On the other hand, they are more transitive than normal intransitive clauses in Nias in that they contain two semantic arguments, neither of which is marked with a preposition, and both of which are marked in the absolutive case, a case otherwise used for arguments of the verb.

Another candidate for semi-transitive status in many languages is clauses containing certain verbs of motion, with which some locative expression is obligatory, as in the examples in (150) from Babungo, which are unacceptable without a locative expression, typically a prepositional phrase.

(150)   a. Ndùlá gɜ̀   táa yìwìŋ
            Ndula go.PERF to   market
            'Ndula has gone to the market'

        b. fəshīa   kò'   fúu tɨ̀
            squirrel climb.PERF on   tree
            'a squirrel climbed on a tree'

But while such clauses may be justifiably classified as semi-transitive, they rarely exhibit any grammatical differences from intransitive clauses with

optional locative expressions apart from the obligatory status of the locative expression, and hence it is in general not necessary to posit a distinct class of semi-transitive clauses, simply on the basis of clauses like these.

The examples above illustrate semi-transitive clauses where this type of clause is determined by specific lexical items. One also finds instances of what one could call semi-transitive clauses in grammatically determined contexts. For example, Yukulta (Keen (1983)), a Tangkic language spoken in northern Australia, generally exhibits ergative case marking, as in (151).

(151)   ṭaŋka-ya=kari   ŋawu   palaṭa
        man-ERG=PRES.TRANS   dog.ABS   hit
        'the man is hitting the dog'

(The final morpheme =*kari* in the first word in (151) is a tense clitic that attaches to the first word in the sentence; this form is used in transitive clauses.) Compare (151) with the corresponding negative sentence in (152), in which the A is in the absolute case and the P is in the dative case.

(152)   walira=ŋka   ṭaŋka-ṛa   ŋawu-ṉṭa   palaṭa
        neg=PRES.INTRANS   man-ABS   dog-DAT   hit
        'the man is not hitting the dog'

This type of semi-transitive clause is parallel to that illustrated above for Djaru in (147), except that here it is grammatically conditioned rather than lexically determined: negative clauses all follow the semi-transitive pattern.

It should be said that semi-transitive clauses are probably not a well-defined cross-linguistic category, in contrast to intransitive clauses and transitive clauses. Rather, in designating these clauses as semi-transitive, nothing is intended beyond observing that they exhibit properties that fall in between those of normal intransitive and transitive clauses. It may well be that in some languages, there is a good analysis of semi-transitive clauses that accounts for their properties, but that the best analysis will vary from language to language. The label 'semi-transitive' indicates nothing more than the fact that the clauses so designated are problematic as far as the traditional distinction between intransitive and transitive clauses is concerned.

2.6   *Clauses with derived verbs*

Our comments in this section will be very brief, since the topic of this section is dealt with in greater detail in other chapters in this anthology. We have restricted discussion of the various types of verbal clauses in this chapter to clauses involving *basic* verbs rather than ones involving some sort of derivation that might result in a clause of a different sort. For example, passive clauses can be thought of as derived in some sense, whether one thinks of it as deriving

passive clauses from active clauses, passive vPs from active vPs, or passive verbs from active verbs. Passive constructions are discussed in chapter 6 by Keenan and Dryer. Other sorts of constructions involving derived verbs include antipassive constructions, noun incorporation, causative constructions, and applicative constructions. A number of these constructions are discussed at length in chapter 7 by Foley.

3   Suggestions for further reading

Two detailed books on nonverbal clauses are Hengeveld (1992) and Stassen (1997). There is an extensive literature on case and grammatical relations that is relevant to the topic of different verbal clause types. Among the basic sources are Blake (1990, 1994), Palmer (1994), and chapter 6 of Van Valin and LaPolla (1997). A number of chapters in this anthology discuss issues related to this, including chapter 3 by Andrews, and chapter 7 by Foley. There is also an extensive literature on ergativity, including Dixon (1994). On split intransitivity, see Merlan (1985) and Mithun (1991).

A number of chapters in *The World Atlas of Language Structure* (Haspelmath *et al.* (2005)) are relevant to topics discussed in this chapter, including Comrie (2005), Siewierska (2005), Haspelmath (2005), and Stassen (2005a, 2005b)

5   Speech act distinctions in grammar

Ekkehard König and Peter Siemund

1   Speech acts and sentence types

In contrast to the traditional view that the function of language is essentially a descriptive one, it is now generally accepted that in speaking we perform actions of various kinds. This is clearest in the case of so-called 'performative utterances' like the following:

(1)   a. I (hereby) order you to leave the room.
         b. I promise you never to be late again.
         c. I hereby declare this meeting closed.
         d. I hereby christen this ship *Queen Elizabeth*.

Sentences like these are special insofar as their utterance in appropriate circumstances amounts to performing the action identified by the finite verb. The typical formal properties of such sentences in English include first person subjects, second person indirect objects, a present tense non-progressive active form of a speech act verb and the deictic adverb *hereby*, but performative sentences may also be in the passive voice, contain modal hedges and a nominalization instead of a verb:

(2)   a. You are cordially invited to come to my birthday party.
         b. I must admit that you have won that argument.
         c. I'll come to see you next week, and that's a promise.

European languages, in particular, have large inventories of such 'performative verbs', but 'performative utterances' of the type in (1) or (2), i.e. sentences whose meaning is such that we can perform the action named by the verb just by saying literally that we are performing it, are rarely used. Most frequently, they seem to occur in institutional settings, where they are part of more elaborate rituals. In fact many, if not all, 'performative verbs' (e.g.

'marry', 'christen', 'appoint', 'resign', 'baptize', 'veto', 'guarantee', and even 'bet' or 'promise', etc.) presuppose the existence of the relevant extra-linguistic social institutions. A second important context seems to be the one where the action performed by an utterance is unclear and needs to be made fully explicit. In most verbal interactions, however, the kind of speech act performed by an utterance is only very weakly determined by the meaning of the sentence uttered. A simple imperative like 'Sit here', for instance could be used as a command, request, offer, advisory, or exhortation, depending on the context, as is shown by the following potential responses: 'Yes, sir' (command), 'Okay' (request), 'No thanks' (offer), 'What a good idea' (advisory), 'Thank you' (exhortation) (cf. Clark (1996:213)). Examples such as these show that it is only the communicative potential of a sentence, a default interpretation, that is determined by its formal and semantic properties. The precise speech act performed by an utterance is the result of an interaction between these properties and various contextual factors, such as the social situation, the current state of an interaction and the background knowledge of speaker and hearer. Moreover, our examples suggest that the precise function (illocutionary force) of an utterance may partly be the result of cooperative negotiations between speaker and hearer.

Three basic sentence types are traditionally distinguished for European languages and have also been found useful for many other languages: declarative, interrogative and imperative sentences.[1] Declarative sentences are primarily and most frequently used for speech acts such as asserting, claiming, stating, but also for accusing, criticizing, promising and guaranteeing. Interestingly enough, all performative sentences are also of the declarative type. Interrogative sentences are typically used for eliciting information, asking questions, introducing deliberations, etc., and imperatives have their basic use in all attempts to get or advise the hearer to do something, i.e. speech acts such as orders, requests, suggestions, prescriptions, appeals, etc. The examples in (3) are instances of the three basic sentence types in English, all involving the same sentence radical and proposition. It is only in combination with other factors, both linguistic and contextual, that such sentences are restricted to expressing specific speech acts.

(3) a. John is taking out the garbage.
     b. Is John taking out the garbage?
     c. Take out the garbage, John.

Any attempt to characterize the patterns and limits of variation across languages in the domain of sentence types cannot simply be based on the assumption that

[1] Note that these three sentences types are also often distinguished in the orthography by different punctuation marks ('.', '?', '!').

the three types mentioned in (3) are distinguished in all languages or exhaust the list of basic sentence types that may be distinguished. What this list shows, however, is that we are looking for grammatical distinctions that can be correlated with a certain use of potential or illocutionary functions. What we need to correlate are, on the one hand, general semantic and/or functional distinctions, as found in various typologies of speech acts (cf. J. Austin (1962); Searle (1976)). Searle's typology, for example, is based on a variety of criss-crossing dimensions (the point or purpose of an utterance, the way the content is related to the world, whether obligations are introduced for the speaker or for the hearer, etc.) and distinguishes five basic types of speech acts: representatives, commissives, declarations, directives and expressives. On the other hand, a cross-linguistically useful definition of sentence types must also be based on formal criteria: the sentence types or, more specifically, the formal properties characterizing those types should ideally form a system of alternative choices that are mutually exclusive, such that each sentence token can be assigned to one type and no sentence token can be a member of more than one type (Sadock and Zwicky (1985:158)). This formal criterion is ideally fulfilled in those cases where the formal markers (inflectional affixes, word order patterns, particles, etc.) identifying the basic sentence types in a language form a system of alternative choices.[2] In Greenlandic Eskimo (Sadock and Zwicky (1985); Sadock (1984)), for instance, the three basic sentence types are identified by different verbal affixes:[3]

(4)   Greenlandic Eskimo
       a. Iga-voq
          cook-DEC.3SG
          'He cooks'

---

[2] Apart from the language material found in the relevant literature, we have based our study on the following language sample (convenience sample): Araona, Chontal (of Oaxaca), Dumi, English, Evenki, Finnish, French, Futunian, Georgian, German, (Modern) Greek, Gulf Arabic, Hayu, Hebrew, Hidatsa, Hua, Hungarian, Indonesian, Italian, Japanese, Korean, Lango, Lezgian, Limbu, Macushi, Malayalam, Mandarin Chinese, Maori, Nama Hottentot, Persian, Punjabi, Rapanui, Russian, Samoan, Shona, Somali, Spanish, Swedish, Tsez, Turkish, Tzotzil, Ute, Wai Wai, Wardaman, Warekena, Welsh, West Greenlandic, Wolof, Zoque.

[3] Additional examples (from Fortescue (1984)) are:

    (i)   angirlasi-   nngil-   anga
          be.homesick   not   1SG.DEC
          'I am not homesick'

    (ii) niri-   riir-   pit
          eat   already   2SG.INT
          'Have you already eaten?'

    (iii) niri-   guk
          eat   2SG.3SG.IMPER
          'Eat it!'
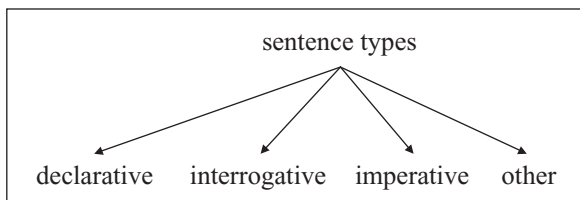
Figure 5.1 Sentence types

b. Iga-va
cook-INT.3SG
'Does he cook?'

c. Iga-git / -guk
cook-IMPER.2SG / -IMP.2/3SG
'Cook (something) / it!'

A similar opposition between three basic sentence types (declarative, interrogative, imperative) is found in Nama Hottentot (Hagman (1977)), and in Khoisan languages in general, but in these languages the markers of sentence type are particles rather than affixes (5). A noteworthy property of this system for marking basic sentence types is that only the particle marking declarative sentences is obligatory. Interrogative and imperative sentences are identified via specific intonation patterns, but can also occur with the relevant particles. The occurrence of a declarative particle in interrogative or imperative sentences is strictly ruled out.[4]

(5)   Nama Hottentot
        a. Declarative: NP + DECL + PredP   (DEC = *ke*, *km*)
        b. Interrogative: NP-à + (INT) + PredP   (INT = *kxa*)
        c. Imperative: NP-à + PredP + (IMPER) (IMPER = *ré*)

In Korean, by contrast, the system of moods (declarative, interrogative, imperative, adhortative) interacts with honorification, i.e. with five or six different verbal affixes being used for each sentence type depending on the speech style (formal, polite, blunt, familiar, intimate, plain) and thus ultimately on the interplay between speaker, hearer and the participants of the situation talked about (cf. Table 5.1, taken from Chang (1996:191)). Evidently, the differentiation is not drawn on all speech levels. These, however, are by no means wide-spread situations. In many languages interrogative sentences can simply be derived from their declarative counterparts through the addition of a particle or tag.

---

[4] The suffix -à in (5b–c) occurs at the right periphery of an NP and has often been interpreted as a case suffix. Since it may mark subjects and objects, as well as oblique NPs, such an analysis is somewhat implausible (Tom Güldemans (p. c.)).

Table 5.1 *Markers of sentence type in relation to speech levels in Korean*

Speech level		Sentence type			
		Declarative	Interrogative	Imperative	Adhortative
High	Formal	(*su*)*pnita*	(*su*)*pnikka*	(*u*)*psio*	(*u*)*psita*
	Polite	(*e*)*yo*	(*e*)*yo*	(*e*)*yo*	(*e*)*yo*
Mid	Blunt	*so*	*so*	*so*	*so*
	Familiar	*ney*	*na*	*key*	*sey*
Low	Intimate	*e*	*e*	*e*	*e*
	Plain	*ta*	(*nu*)*nya*	*la*	*ca*

In Shona (Bantu), for instance, it is the addition of the particle *hèré* (yes/no questions) or of an interrogative pronoun (constituent questions) that turns a declarative into an interrogative sentence:

(6)   Shona
     a. ndì-nó-tàùr-à
        1SG-PRES-speak-DECL
        'I speak'

     b. ndì-nó-tàùr-à   hèré?
        1SG-PRES-speak-DECL INT
        'Do I speak?'

And in Japanese, declarative sentences are differentiated from imperatives by a clear paradigmatic contrast, but are unmarked with respect to interrogatives:

(7)   Japanese
     a. Sakana-o tabe-ro/-te (kudasai)!
        fish-ACC   eat-IMPER
        'Eat the fish!'

     b. Sakana-o tabe-ru.
        fish-ACC   eat-PRES.DECL
        'I eat fish'

     c. Sakana-o tabe-ru-ka?
        fish-ACC   eat-PRES.DECL-INT
        'Do you eat fish?'

Another common difficulty with the view of a clear paradigmatic opposition between the three basic sentence types under discussion is the fact that the imperative is often expressed by a specific inflectional form even in languages which do not distinguish the two other types by morphological means. In such languages the imperative is often analysed as being one option in a system of
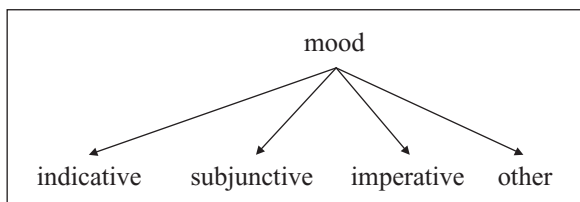
Figure 5.2 Mood distinctions

'mood', which also includes the categories 'indicative', 'subjunctive', 'conditional', 'optative' and perhaps others (Palmer (1986:23ff.)).

A clear paradigmatic opposition between basic sentence types is sometimes found with embedded sentences, particularly in European languages. In such structures the relevant opposition is expressed by different complementizers or the lack thereof (8).

(8) a. I knew that John did it. (declarative)
     b. Fred wonders whether/if John did it. (interrogative)
     c. Fred asked John to help him. (jussive)

Note, however, that imperatives cannot be embedded and that there is thus no 'imperative complementizer'. Such complementizers have no bearing on the speech acts that may be performed by the sentence containing them, unless they occur in an independent, non-embedded sentence (9).[5]

(9) a. That I should live to see this!

     b. German
        A: Bist Du müde? – B: Und ob   (ich müde bin)
           are you tired   and whether I   tired am
        'A: Are you tired? – B: Am I ever!'

The set of expressive devices used in the languages of the world for the differentiation of basic sentence types includes all those that are generally considered to be part of grammatical systems: intonation, inflection (more specifically the addition or omission of inflectional affixes), word order, and the addition, omission, or substitution of constituents.[6] Intonation clearly has a special role, since it is never in paradigmatic contrast with another strategy. Moreover, any

---

[5] Some languages (e.g. German) exploit this use of complementizers in non-embedded clauses quite systematically for the characterization of minor sentence types.

[6] Croft (1994:462ff.) argues that the formal properties distinguishing the basic sentence types are particularly salient ones, even from a cross-linguistic perspective. Moreover, he proposes that the 'distance' between sentence types can be iconically motivated by taking into account their structural differences: the more substantial the formal differences, the farther the sentence types are apart in terms of their use potential or illocutionary force.

intonation contour, or at least most of them, can apparently be superimposed on any segmental structure, an option that is frequently exploited in so-called 'echo-questions'. It is for these reasons that we hesitate to regard intonation as the most important strategy for the identification of sentence types and in fact hesitate to identify a formal sentence type in a language purely in terms of intonation.

In spite of the difficulties of basing the distinction between sentence types on formal paradigmatic oppositions, it is still possible to draw such distinctions for most languages. Despite a certain heterogeneity in the formal inventory used to identify different sentence types, sentences can usually be assigned to one and only one basic sentence type within a language without fulfilling additional conditions. The labels 'declarative', 'interrogative' and 'imperative' can then be assigned to these formal types on the basis of their typical use.[7] Moreover, these three basic sentence types – which have played a prominent role in the analysis of European languages – also seem to be more clearly identifiable across languages than other 'basic' types mentioned in the descriptions of various languages. More often than not, these other basic types are simply the result of an interaction of two formal properties, rather than being marked by a single grammatical device: 'prohibitives' ('Don't do that') can often be analysed as negative imperatives; 'optatives' ('(May) God bless you') may simply be based on the third person subjunctive; 'hortatives' ('Let's go') may be further analysable as first person plural imperatives; and exclamatives are typically based on interrogatives or declaratives (*Isn't she wonderful!*; *He is such a nuisance!*). Of course, all of these categories may find a unique expression in a language – and it is for those cases that labels like 'prohibitive', 'optative', 'exclamative', etc. should be reserved – but they may also be the combinatorial result of several features. The three basic types, declarative, interrogative and imperative, by contrast, seem to be clearly identifiable on the basis of only one formal property in the vast majority of languages, even if the relevant properties do not always constitute a neat paradigm of formal oppositions.[8] What the languages and data examined in this paper do not support, however, is the view that three or more sentence types can be differentiated and identified on

---

[7] There are cases, however, where such an identification does not seem to be possible unless and until a second property (e.g. intonation) is considered. Polite imperatives in German are a case in point. In these imperatives the distal (polite) form of address (*Sie*) has to be used as subject and thus the resultant sentences have the formal properties of interrogative structures. A sentence like *Kommen Sie mit* (lit.: 'Come you along') would still be classified as an interrogative if it has a rising rather than a falling intonation contour. Similarly, verb-first structures such as *Kommt ein Mann zur Tür herein* (lit.: 'Comes a man through the door') can be used both as interrogative sentences and as declaratives. Such declaratives, however, require specific contextual conditions for their use (cf. Önnerfors (1997)).

[8] We do not wish to rule out languages in which more than one formal property is necessary for marking basic sentence types. Although such languages seem to be rare, the truth is that little systematic work has been done on this matter.

the basis of formal properties in all languages. Allowance must be made for cases of overlap between two sentence types in specific contexts, as well as for cases where there is no clear formal distinction between two types at all, typically between declaratives and interrogatives. Moreover, a clear assignment of sentence tokens to a specific type is often only possible on the basis of more than one criterion.

There is a well-known problem for the view that the basic sentence types have a certain functional potential or default interpretation that needs to be addressed at this point. So far we have assumed that declarative sentences have an essentially descriptive function, that imperatives are primarily used for directive speech acts and that interrogative sentences are primarily used as questions. Such functional criteria are in fact essential for an identification of sentence types as declarative rather than interrogative across languages. Unfortunately, such a view of the connection between form and illocutionary function is clearly an oversimplification:

(10)   a. It is terribly cold in this room.
       b. Could you please close the window?
       c. Who likes being criticized?
       d. You must be feeling very tired.
       e. Why don't you buy some stocks?
       f. Waiter, what's that fly doing in my soup?

The examples in (10) show that in English both a declarative sentence (10a) and an interrogative sentence (10b) can be used to perform directive speech acts, i.e. as requests for the addressee to close the window. Together with a falling intonation contour and an accent on the participle, the example in (10c) would typically be used as a rhetorical question, i.e. as a statement expressed by an interrogative sentence. A sentence like (10d) could be intended as a request for information (question) expressed in declarative form and sentences like (10e) are more likely to be used as suggestions than as requests to supply a justification for a certain course of action. The only use imaginable for (10f) is that of an act of criticizing.

Several ways of dealing with such asymmetries have been discussed in the literature (see Levinson (1983:263ff.)). The most extreme position is simply to abandon the assumption of clear form–function correlations. In this view, illocutionary force is purely a matter of context and has no direct and simple correlation with sentence form and sentence meaning. But such a theory clearly misses very important and obvious generalizations, as will be shown in the body of this chapter. Another theory ('inference theory'), which avoids such undesirable consequences, is based on the assumption that sentences like (10a–f) are indeed used as statements and questions, respectively, but in addition to having this illocutionary force they also have the force of a request

(10a, b), a statement ('Nobody likes being criticized'), a question ('Are you tired?') and a suggestion ('You should buy some stocks'). This secondary function is indirectly expressed – hence the term 'indirect speech acts' – and based on conversational inferences and principles of conversational interaction which ultimately trigger such inferences. Such inferences rely heavily on contextual factors. A statement of discomfort such as (10a) or *I need something to drink* can only be used as an indirect request if there is a hearer present who feels under an obligation to serve the interests of the speaker. In other words, such hints could simply be overlooked.

This inference theory does not only solve the problems raised by examples such as (10), but also accounts for a variety of interesting facts and generalizations, only two of which will be mentioned here: typical responses to requests expressed by interrogative sentences like (10b) include both a verbal reaction ('yes') and a non-verbal one (carrying out the action) and thus support the view that two actions are performed by such indirect speech acts. Furthermore, such comparative studies as have been carried out in this domain show that the sentences that can be used to perform speech acts indirectly in English have clear parallels in other languages and that therefore the principles for using sentences in indirect speech acts must be very similar for different languages and cultures. What may differ from language to language is the addition of certain expressions to such sentences which would normally be used for the direct performance of the relevant speech act. In German, for instance, assertive particles like *schon* are not only used in declarative sentences, but also in rhetorical questions. The German counterpart of (10c) may thus take the following shape:

(11)   German
        Wer   wird schon   gern   kritisiert?
        who   is   ASSERTIVE   voluntarily criticized
        'Who likes being criticized?'

The addition of *schon*, however, clearly indicates that the relevant sentence is used as a statement, rather than as a question, just as the addition of *please* to an interrogative sentence indicates that the sentence is used as a request (*Could you please get me some coffee?*). As a result such patterns may become conventionalized for a certain use and a new sentence type can be created.

2   Declarative sentences

Declarative sentences are conventionally and typically used to perform representative (descriptive) speech acts such as assertions, reports, acts of complaining and bragging, but also acts of predicting and promising. All such acts convey the belief of the speaker that the proposition expressed is true or will turn out to be true. Of course, declarative sentences can also be used for directive speech acts, such as requests (12a) and commands (12b):

(12)   a. Swedish
           Du   räcker   mig   salt-et
           you   pass   me   salt-DEF
           'Can you pass me the salt?'
           (literally 'You pass me the salt!')

        b. German
           Du   bezahlst   jetzt   sofort   deine   Schulden!
           you   pay.2SG   now   right.now   your   debts
           'You are going to pay your debts right now'

As is shown by the preceding examples, however, a number of additional con-
ditions have to be met for such a use to be possible: only in combination with
second-person subjects, with a non-past tense and with a non-stative predicate
can a declarative sentence be used as a request or as an order.

2.1   *Declaratives in relation to the other basic types*

Among the three major types distinguished so far, declarative sentences are the
'unmarked' member, in several senses of this term. They can be considered
as unmarked relative to interrogative and imperative sentences for at least the
following reasons:
  (i) Declarative sentences are the most frequent sentence type.
 (ii) The word order exhibited by declarative sentences (SOV, SVO, VSO, etc.) is
      normally regarded as the basic word order of a language.
(iii) Declarative sentences are less restricted in their distribution compared
      to the other two types. Embedded clauses often have the same formal
      properties as declarative sentences and, as pointed out above, 'performative
      utterances' are also of the declarative type.
(iv) Declarative sentences exhibit the full paradigm of tense–aspect combina-
      tions available in a language, in contrast to imperatives.
 (v) Declarative sentences may be used to express most of the speech acts dis-
      tinguished in the typology by Searle, i.e. they can be used as representative
      speech acts (*It is raining*), as commissives (*I will never again forget your
      birthday*), as directives (*You know what you have to do*), as expressives (*I
      am sorry*) and as declarations (*He is guilty*).
(vi) In many languages, interrogative sentences and partly also imperatives can
      be analysed as being the result of some operations (adjunction, omission,
      change of word order) performed on declaratives, rather than the other way
      round.
   In languages that have systems of inflectional mood distinctions, declaratives
are characterized by the indicative mood, of course, but so are interrogatives
and sometimes even conditionals. So, perhaps, a falling intonation contour is
a more reliable formal indicator of declaratives, but such a claim can only be

made for the tiny subset among the world's languages whose prosodic properties have been examined in sufficient detail. Overall, the view that declaratives do not have a specific formal marker is wrong: declarative sentences are characterized by precisely those formal properties which distinguish them from other sentence types, e.g. the absence of an interrogative marker and the absence of an imperative inflectional form, or the presence of a specific finite verbal form.

As already mentioned, interrogatives and partly also imperatives can frequently be described as being the result of modifying declaratives in some way; by changing the word order (inversion), adding a particle, etc. English interrogatives, for example, exhibit inversion of subject and auxiliary verb relative to declarative word order, unless the interrogative pronoun is the subject:

(13)   a. You are taking the train.
         b. Are you taking the train.
         c. Who did you see?
         d. Who saw the thief?

Another typical situation is the formation of interrogative sentences from declarative sentences by adding a special particle. Such interrogative particles are quite frequent among the languages of the world. They are used in Tzotzil, a Mayan language spoken in Mexico (14), and also in French (15):

(14)   Tzotzil
         a. ch-a-bat
             ASP-2SG-go
             'You are going.'

         b. mi   cha-a-bat
             INT ASP-2SG-go
             'Are you going?'

(15)   French
         a. Jean est malade.
             John is sick
             'Jean is sick.'

         b. Est-ce que Jean est malade?
             'Is John sick?'

In those languages in which the formal means for marking the basic sentence types form a system of paradigmatic oppositions, declarative sentences receive special marking.[9] However, our overall impression is that languages with such

---

[9] For oppositions of this kind the term 'marked declaratives' has been proposed (cf. Sadock and Zwicky (1985)). We will avoid this terminology since, as stated above, also so-called 'unmarked declaratives' have specific formal properties which distinguish them from other sentence types.

paradigmatic oppositions are relatively rare. We have found paradigmatic oppositions of three different kinds. Firstly, the three major sentence types can be distinguished inflectionally within an elaborate mood system. In addition to Eskimo (cf. (4) above), there are some language families in Brazil (Tariana, Tucano, Jarawara, etc.) with such mood systems, which typically include several imperatives. Secondly, there may be different particles for declarative, interrogative, and imperative sentences, which are mutually exclusive. The only language family in our sample illustrating such a paradigm is Khoisan (cf. (5) above). Thirdly, the three basic types of sentences may be assigned different word order patterns. In German, for example, the basic word order is generally assumed to be SOV, the word order manifested by subordinate clauses (cf. (16)). From this basic word order the sequential organization of the three sentence types can be derived as follows: (i) interrogatives are the result of moving the finite verb to initial position (17a);[10] (ii) declarative main clauses involve the additional operation of moving a constituent (the topic) in front of the finite verb, thus resulting in the order TVX, i.e. topic-verb-the rest (17b); (iii) imperatives exhibit the same order as declaratives, but typically lack a (second person) subject and are also characterized by a special (impoverished) inflection (17c):

(16)   German
       Der Kanzler   behauptet [dass ein Freund ihm das Geld   gab]
       the chancellor claims   that a   friend him the money gave
       'The Chancellor claims that a friend gave him the money.'

(17)   German
       a. Gab ein Freund ihm das Geld?
          gave a   friend him the money
          'Did a friend give him the money?'

       b. Ein Freund gab ihm das Geld.
          a   friend gave him the money
          'A friend gave him the money.'

       c. Gib ihm das Geld!
          give him the money
          'Give him the money!'

The differentiation of the basic sentence types through word order is a typical feature of the Germanic languages (German, English, Norwegian, Swedish, Danish) and rarely found outside of Europe.[11]

---

[10] Under specific circumstances verb-first structures can also be used as declarative sentences, as exclamative sentences and as conditional antecedents.

[11] This picture is not fully adequate for German and can only be justified in terms of frequency. In the overwhelming majority of cases, main clause declaratives are indeed verb-second structures,

2.2   *Interaction with evidentiality*

In languages with markers signalling the kind of evidence (hearsay, common knowledge, first-hand visual evidence, etc.) on which a claim is based or the degree of strength with which an assertion can be made, these so-called 'evidential markers' can normally only be combined with declarative sentences. This is scarcely surprising given that declarative sentences are typically used to express claims, assertions, statements about the world (of discourse) and thus indicate an attitude of belief (in the truth of the proposition expressed). However, this co-occurrence restriction is not sufficient proof for saying that evidentials are markers of declarative sentences.

A language that is frequently mentioned in this context is Hidatsa (a Northern Plains Siouan language), which has five different particles (*ski*, *c*, *wareac*, *rahe*, and *toak*), expressing different kinds of evidential meaning. These morphemes occur at the end of a sentence, but do not co-occur with one another nor with the particles marking interrogatives or imperatives (G. H. Matthews (1965:99f.); Palmer (1986:70); Zaefferer (1990:222)). The names and glosses Matthews provides for these evidential particles are given in Table 5.2

A similar system can be found in Tuyuca (Brazil and Colombia) where five different types of evidential meaning are formally distinguished: (a) visual evidence; (b) non-visual (perceptual) evidence; (c) apparent/inferential; (d) secondhand/quotative; (e) assumed (Barnes (1984:257); Palmer (1986:67)). The following examples provide illustration for these distinctions. Each sentence would be translated into English as 'he played soccer', but is based on different evidence, as is indicated in brackets:

Table 5.2 *Evidentials of Hidatsa*

Evidential marker	Label	Description
ski	emphatic	definitive knowledge
c	period	believed, desired or felt by the speaker
wareac	quotative	general knowledge
rahe	report	learnt from hearsay
toak	indefinite	something the speaker does not know and thinks the listener does not either

but verb-first structures can also be used as declaratives under specific conditions (cf. Önnerfors (1997)). Lohnstein (2000) has shown that it is the interaction between word order and verbal mood that differentiates between sentences types in German clearly and unambiguously.

(18)   Tuyuca
a. díiga apé-wi   (I saw him play.)
b. díiga apé-ti   (I heard the game and him.)
c. díiga apé-yi   (I have seen circumstantial evidence that he played.)
d. díiga apé-yigɨ   (I obtained the information that he played from s.o. else.)
e. díiga apé-hīyi   (It is reasonable to assume that he played.)

<div align="right">Palmer (1986:67)</div>

Finally, in Jaqaru (Jaqi, South America) and related languages, the verbal suf-
fixes encoding evidential meaning stand in paradigmatic opposition to inter-
rogative suffixes, may occur only once per sentence and thus appear to form a
paradigm of contrastive sentence types:

(19)   Jaqaru
       a. Amrucha-txi. (polar interrogative)
         'Is X well?'

       b. Amrucha-wa. (declarative, personal knowledge)
         'X is well.'

       c. Amrucha-mna. (declarative, knowledge through language)
         'X is well, they say.'   Hardman (1986:129f.)

In some languages evidential markers combine with all tenses. Quite frequently,
however, they are restricted to certain tenses, typically the past or perfect
tense. In Turkish, for instance, there is an obligatory choice for all past tense
expressions: either the suffix –*dI* (marker of direct experience) is chosen or the
marker of indirect experience (inference, hearsay) –*mIş*:

(20)   Turkish
       Ahmet gel-miş.
       Ahmet come-QUOT/INFER
       'Ahmet came / must have come.'

And in Tsez (Daghestanian, Northeast Caucasian) there is a morphologically
marked distinction between witnessed (-*s*(*i*)) and unwitnessed (-*n*(*o*)) events in
the past (cf. Comrie and Polinsky (in press)):

(21)   Tsez
       a. Kid   mek'u-n.
         girl.ABS be.hungry-PAST.UNWITNESSED

       b. Kid   mek'u-s
         girl.ABS be.hungry-PAST.WITNESSED
         'The girl was hungry'

Languages like English and German which do not have such evidential markers
either do not indicate the type of evidence their claim is based on at all or they

express the relevant meaning by a higher clause (22), by a modal verb (23) or by a modal particle (24):

(22)   I see/hear/understand you are leaving this country.

(23)   German
   a. Karl   *soll*   mit   dem   Papst   gesprochen   haben.
      Charles   shall   with   the   Pope   spoken   have
      'Charles is said to have talked to the Pope.'

   b. Karl   *will*   mit   dem   Papst   gesprochen   haben.
      Charles   *will*   with   the   Pope   spoken   have
      'Charles claims to have talked to the Pope.'

(24)   German
      Du   blutest   *ja*! (visual evidence)
      you   bleed   PRT
      'Why, you are bleeding.'

The preceding examples of a close interaction between declaratives and evidentiality show that there is no generally accepted inventory of categories and terms that can be used unambiguously for a cross-linguistic description in this domain. Certain descriptive labels, such as 'visual evidence', 'auditory evidence', 'sensory evidence' / 'non-visual', 'hearsay', 'quotative', 'inferential', 'personal knowledge', 'general knowledge', 'reportive', '(un)witnessed', 'dubitative', etc., are found quite frequently in individual descriptions and descriptive surveys (cf. Palmer (1986); Chafe and Nichols (1986)), but it is not clear which of these terms describe similar, different or exactly the same phenomena. Nor is it clear what the maximal number of possible distinctions is, which categories are superordinate to others and whether all of the distinctions drawn in a language can be ordered on a single dimension. Mood oppositions such as the ones found in Hidatsa (see Table 5.2) suggest that two different subsystems should be distinguished (Palmer (1986:53)): judgements (strength of the speaker's commitment) and evidentials (type of evidence, channel/source of information). A further interesting problem that should be mentioned at this point is the relationship between an affix or particle expressing ignorance or doubt ('dubitative') and interrogative markers. In a variety of languages the same suffix or particle is used for both to indicate that the speaker does not know whether a sentence is true, and may thus characterize the sentence as either declarative and dubitative or as interrogative (cf. Sadock and Zwicky (1985:169f.)).

3   Interrogative sentences

The fact that most, if not all, languages have sentences of special structural types for asking questions clearly demonstrates how central this activity is to human

communication. These 'interrogative sentences' are conventionally associated with the speech act of requesting information. Interrogative sentences fall into two major classes depending on their syntactic and semantic properties. It is obvious that the two interrogative sentences given in (25) below have different syntactic structures and are typically used for different types of requests or inquiries.

(25)   a. Do you believe in miracles?
        b. Who discovered America?

Sentence (25a) is an instance of a so-called 'polar interrogative sentence' or 'yes/no' question', where, quite in contrast to (25b), the expected answer simply consists in providing a truth value for the corresponding declarative sentence. Polar interrogatives are typically used to inquire about the truth or falsity of the proposition they express. It should be borne in mind, however, that answers to polar questions can plausibly assume any value on a scale between 'true' and 'false', as, for example, 'perhaps', 'possibly', 'quite likely', etc. 'Constituent interrogatives' of the type exemplified by (25b), also known by the name 'information questions', receive answers that provide the kind of information specified by the interrogative word (*wh*-words like *who*, *when*, *how* in English) contained in it, i.e. some expression denoting a human being in the case of (25b). Semantically, constituent questions can be analysed as open propositions with interrogative words signalling the relevant variable positions. As example (26) demonstrates, constituent interrogatives may contain more than one interrogative word:

(26)   Who did what to whom?

Strictly speaking, a clear distinction has to be made between interrogatives as a type of sentence or clause and their semantic counterpart, i.e. questions (cf. Huddleston (1994)). Such a distinction is not consistently found in the literature, however.

    The formal differentiation between polar interrogatives and constituent interrogatives is a stable cross-linguistic parameter and for each type of interrogative sentence a limited number of recurring, and sometimes overlapping, coding strategies can be identified across the world's languages (Siemund (2001)). With respect to its meaning, another type of question can be distinguished:

(27)   Do you prefer beer or wine?

In the prototypical case, such 'alternative questions' are used to ask the addressee to decide which of two or more alternatives holds, i.e. is true or not. Apparently, alternative questions have a lot in common with polar questions, but

strictly speaking the answer set is different.[12] Be that as it may, for the purposes of the present study alternative questions can be neglected since, at least from our current perspective, they do not seem to show any striking typological variation. Nevertheless, alternative questions will become important in the subsequent discussion of coding strategies for polar questions.

3.1   Polar interrogatives

Essentially six ways of expressing polar questions are encountered across the languages of the world, by and large independently of their genetic affiliation. In decreasing order of their frequency of occurrence these are: (i) special intonation patterns, (ii) interrogative particles, (iii) the addition of special tags, (iv) disjunctive-negative structures, (v) a change in the relative order of constituents and (vi) particular verbal inflection. By far the most prominent among these six strategies are the use of a special intonation pattern and the addition of a particular particle. Interrogative tags are also quite frequent, but typically add some bias with respect to the answer expected. Disjunctive-negative constructions of the type 'A or not A' are found in some Asian languages. A constituent order different from the one used for declarative sentences is mainly restricted to Indo-European languages. On the whole, these strategies are much less frequent in comparison to the two mentioned at the beginning. The encoding of interrogativity by verbal inflection, i.e. an interrogative mood, is extremely rare and mainly restricted to polysynthetic languages. Many languages have more than one strategy – although one of these is usually primary – and typically use intonation in combination with one of the remaining five strategies.

3.1.1   Intonational marking   Broadly speaking, the intonation contour used in interrogative sentences is the opposite of the one found in declaratives. While it is typical of declaratives to show falling intonation, the great majority of languages use rising intonation in conjunction with interrogatives. Exceptions to this generalization are very rare indeed. However, cases of rising intonation in declaratives and a falling contour in interrogatives do occur and are, for example, reported from Fanti (Niger-Congo, Kwa) and Grebo (Niger-Congo, Kru). The reason for the predominance of rising intonation in interrogatives is usually seen in the fact that high pitch signals uncertainty, indecision, hesitation and also insecurity. Low pitch, by contrast, is assumed to convey confidence, assurance and certainty (cf. Ohala (1983, 1994)). The intonation contours of declaratives and interrogatives thus provide good illustration for the principle of iconic motivation. The Italian example in (28) illustrates the standard pattern.

---

[12] As a matter of fact, according to their answer set alternative questions are like constituent questions.

(28)   Italian
         a. Suo marito   è ancora \malato. (statement)
            her husband is still   ill
            'Her husband is still ill.'

         b. Suo marito   è ancora /malato? (question)
            her husband is still   ill
            'Is her husband still ill?'

Of course, the actual shape of the rising contour is not the same in all languages. There is a clear tendency for languages to place the rise toward the end of the contour, but interrogatives marked by an initial rise do occur also, albeit somewhat sporadically. Moreover, Ultan (1978) reports a number of recurrent shapes which final rises apparently can assume, the most important of these being higher ultima (Vietnamese), higher penult (Chontal, a Hokan language), higher pitch on last stressed vowel (Bashkir, a Uralian language of the Altaic group) and rising toward last stressed vowel (Hebrew, Semitic).

   Even though declarative sentences with rising intonation contours can often be analysed as expressing questions, it is highly doubtful whether they should be regarded as instances of the form type 'interrogative' (cf. Huddleston (1994:428)). Several facts argue against such an analysis. First of all, the domain of the rising contour can be wider than that of at least some of the other signals used for the interrogative type and comprise a coordination of clauses (29a). On the other hand, the superordinate clause of a complex sentence may be outside the scope of the question proper (29b).

(29)   a. So, Kim went to the meeting but you stayed at home?
         b. I don't suppose Jack will contribute to our cause?

Secondly, declarative sentences with rising intonation do not license negative polarity items like *ever*, *any*, *at all*, etc., in contrast to clear cases of interrogative sentences:

(30)   a. Have you ever met him?
         b. *You have ever met him?

Thirdly, declarative sentences with rising intonation, like (31) in English, cannot be used as neutral questions in many languages:

(31)   a. You stole the money?
         b. He believes in God?

Like other declarative sentences, such sentences generally express a commitment. In contrast to declaratives with falling intonation, however, it is a commitment of the addressee and thus an assumption about the answer that is expressed.

Therefore, such sentences cannot be used as an opening move in a zero context, i.e. in a context where no assumptions about the answer are justified.

Finally, rising intonation can be combined with any form type and it has been shown that rising intonation plays a fairly minor role in signalling questions with either inverted or uninverted structures in English (Geluykens (1988)). If criteria like these have cross-linguistic relevance, the conclusion seems inescapable that languages which use declaratives with rising intonation to express questions do not have the sentence type 'interrogative'. On the other hand, there are languages where 'declarative sentences' with rising contours meet at least some of the criteria generally considered relevant for the identification of genuine interrogatives (licensing of negative polarity items, use in neutral contexts, etc.). In Russian, for instance, sentences without the interrogative marker *li* but with rising intonation meet the relevant criteria so that their categorization as 'interrogatives' seems justified (Meyer and Zybatow (2003)).

3.1.2   Interrogative particles   The addition of interrogative particles to declarative sentences is another way of deriving interrogatives. Their precise position is subject to considerable typological variation, but to have interrogative particles in sentence-final position seems the most widely used option. Interrogative particles also occur in constituent interrogatives, but mostly optionally so. The contrast between declarative and interrogative exemplified by the pair of Japanese sentences in (32) is representative of the situation found in many languages:

(32)   Japanese
      a. Yamada-san wa   ginkoo de hataraite-imasu.
         Yamada-Mr TOP bank   at working
         'Mr Yamada works at the bank.'

      b. Yamada-san wa   ginkoo de hataraite-imasu ka?
         Yamada-Mr TOP bank   at working   INT
         'Does Mr Yamada work at the bank?'   Hinds (1984:158)

Further examples of interrogative particles include French *est-ce que*, Polish *czy*, Finnish *ko/kö*, Mandarin *ma*, Slavic *li*, Turkish *mi*, Indonesian *kah*, Bengali *ki*, Kannada *e:nu*, etc. In Bengali and Kannada the interrogative particle is homonymous with the interrogative word for 'what'. The case of a sentence-initial particle is illustrated by the Persian example in (33), (see Mahootian (1997:9)). Another such language is Tzotzil (see (14b) above).

(33)   Persian
       Aya in   gorbe-ye šoma-st?
       INT this cat-LINK you-is
       'Is this your cat?'

For Russian, Turkish and Ute (Uto-Aztecan), but also Latin and Finnish, the relevant interrogative particles (*li*, *mi*, *aa*, *ne*, *kol kö* respectively) should probably be categorized as affixes or at least clitics, because in these languages they are not construed with the entire sentence, but are always attached to a particular constituent. Givón (1984b:219f.) shows for Ute *-aa* that it always occurs after the first constituent and that it is in effect enclitic to it (34).

(34)   Ute
    a. mamá-ci-aa   'u   wų́ų̩ka-pų̩gá?
       woman-SUBJ-INT that.SUBJ work-REM
       'Did the woman work?'

    b. kų́aw-aa   páĝa-kway-kya?
       yesterday-INT leave-go-ANT
       'Did (she) leave yesterday?'

Turkish *mi* usually occurs in sentence-final position immediately after the predicate. It takes scope over the entire sentence and there are good reasons to assume that it cliticizes onto the predicate (it shows vowel harmony with the stem and is never stressed although word stress in Turkish is word-final). Nevertheless, the Turkish interrogative particle may also attach to constituents within a sentence. This can be observed in focussing constructions, where the scope of the particle is restricted to the relevant constituent (Kornfilt (1997:191)):

(35)   Turkish
       kitab-ı   Hasán mı Ali-ye   ver-di?
       book-ACC Hasan INT Ali-DAT give-PAST
       'Did HASAN give the book to Ali?'

Given that interrogative particles preferably occur adjacent to the predicate, at least in the unmarked case, it should be possible to make certain predictions concerning the position of interrogative particles and the basic word order pattern of a language. A plausible prediction, and this is by and large borne out by the data, would be that verb-final languages mostly have sentence-final particles whereas verb-initial languages tend to have sentence-initial particles (Greenberg (1966:81)). Nevertheless, there are also some counter-examples to this generalization, as the examples from Persian (33) and Ute (34) show (both SOV). Moreover, apart from general concerns about drawing a distinction between basic word order types, another complicating factor for any such generalization is that languages may have more than one interrogative particle which may differ in its distribution. There seem to be no preferences for the position of interrogative particles in SVO languages. Thai and Yoruba, for instance, have final particles whereas the particle in Lithuanian is initial, but all belong to the basic type SVO.

Korean is one of those languages which have more than one interrogative particle, although they all occur at the end of a sentence. The factor governing their distribution is the level of formality (honorification): high style (*su*)*pnikka*/*(e)yo*, mid style *so*/*na*, low style *el*/(*nu*)*nya* (cf. also Table 5.1). The position indicated by 'X' in (36) below can be occupied by any of these expressions (Chang (1996:84)).

(36)   Korean
        Kui-nun cal   cwumwusi-X
        he-TOP well sleep-INT
        'Does he sleep well?'

Interrogative particles are often closely related to expressions introducing conditional subclauses. As is shown by Russian *esli*, conditional markers can develop out of interrogative particles (*est'* + *li*) and, in fact, consistently do so in language after language (see Traugott (1985:291)). In Hua (a Papuan language, cf. Haiman (1978:570–1)), interrogative particle and conditional marker are formally identical:

(37)   Hua
        a. E   -si   -ve baigu   -e
           come 3SG.FUT INT will.stay 1SG
           'If he comes, I will stay.'

        b. Fri -si   -ve
           die 3SG.FUT INT
           'Will she die?'

3.1.3   Interrogative tags   Closely related to interrogative particles are interrogative tags, the main difference between the two kinds of expressions being that tags, apart from characterizing sentences as questions, also contribute a certain bias by raising expectations toward either a positive or a negative answer. This is illustrated by the English examples in (38), where the (a) sentence expects a positive answer, but the (b) sentence one that is negative:[13]

(38)   a. You like ice-cream, don't you?
        b. You don't like ice-cream, do you?

Although it is a reasonable approximation to the facts to say that negative tags presuppose a positive answer and vice versa, Ultan (1978) points out that the answer induced by a tag question depends to a greater extent on the polarity of the declarative sentence used for forming the interrogative and less so on the

---

[13] Note that such bias may also be produced by polarity items such as *yet* and *already*:

        Have you not eaten yet?
        Haven't you eaten already?

polarity of the tag. Tag interrogatives of type (38a) are clearly the most frequent pattern, but the second most frequent is a combination of affirmative sentence, affirmative tag and positive answer. Cases like (38b) belong only to the third most frequent type.

Another important difference between interrogative tags and interrogative particles is that tags almost exclusively occur at the end of a sentence, quite independently of the basic word order pattern. One exception to this generalization is the Persian tag *mæge* (39). Moreover, tags are formally not particles, but occur as either (content) words (40), phrases (41) or clauses (42).

(39)   Persian
        mæge un mašin-e to nist?
        INT   that car-LINK you isn't
        'Isn't that car yours?'

(40)   Russian
        Ty ego slyšal, pravda?
        you him heard true
        'You heard him, didn't you?'

(41)   German
        Er ist sehr reich, nicht wahr?
        he is very rich   not   true
        'He is very rich, isn't he?'

(42)   Turkish
        Ahmet dün   sinema-ya gi-ti,   değil   mi?
        Ahmet yesterday cinema.DAT go-PAST, NEG.COP INT
        'Ahmet went to the movies yesterday, didn't he?'

                                              (literally 'isn't it so?')

3.1.4   Disjunctive-negative structures   A completely different strategy of forming polar interrogatives from the ones discussed so far is found in Mandarin Chinese and some other Asian languages, as well as in certain languages spoken in Papua New Guinea (Amele, Kobon). What is remarkable about the relevant constructions is their affinity to alternative interrogatives. The example given in (43) looks like an alternative interrogative in which the conjunction is missing and where the second conjunct is the negation of the first. This type of polar interrogative is also known as the 'A-not-A construction' (C.N. Li and Thompson (1984)).

(43)   Mandarin Chinese
        tā   zài jiā   bu zài jiā?
        3SG at home NEG at home
        'Is s/he at home?'

The complete structure as shown in (43), however, is not used very frequently and usually replaced by either of the reduced constructions given in (44).

(44)   Mandarin Chinese
        a. tā zài bu zài jiā?
        b. tā zài jiā bu zài?

What is noteworthy about Mandarin Chinese is that the very same construction, when appended to a declarative sentence, serves the function of an interrogative tag. In other words, the resulting questions are not neutral any longer with respect to their expected answer. Nevertheless, the set of A-not-A structures used as tags is extremely limited and highly lexicalized. Apart from *duì bu duì* 'right not right', as shown in (45), there are only *hǎo bu hǎo* 'good not good', *shì bu shì* 'is not is', *kěyǐ bu kěyǐ* 'may not may', and *xíng bu xíng* 'okay not okay' that occur with significant frequency.

(45)   Mandarin Chinese
        zhāng-sān xǐhuan hē   jiǔ,   duì bu   duì?
        Zhang-san like   drink wine right NEG right
        'Zhang-san likes to drink wine, right?'

3.1.5   Change in the order of constituents   Mostly restricted to Indo-European languages and predominantly to the Germanic branch, by contrast, is the use of a special word order for polar interrogatives. Outside this genetic group, this strategy is very rare. Two non-Indo-European languages known to us that also make use of it are Finnish and Malay. The most common case is to put the finite verb into sentence-initial position while retaining the relative order of the other constituents (46–49):[14]

(46)   Swedish
        a. Lars   läser tidningen.
           Lars   reads the.newspaper
           'Lars is reading the newspaper.'

        b. Läser Lars tidningen?
           reads Lars the.newspaper
           'Is Lars reading the newspaper?'

(47)   a. She is a translator.
        b. Is she a translator?

---

[14] As is shown by our examples in (47–49), in English (with the exception of *be* as a main verb and, for some speakers, *have*) only auxiliary verbs can be shifted to a position preceding the subject. If the corresponding declarative sentence does not contain an auxiliary, the all-purpose auxiliary *do* is introduced. This *do* is fronted and carries the tense.

(48) a. She can swim.
      b. Can she swim?

(49) a. He knows a translator.
      b. Does he know a translator?

From this it follows that inversion of the verb-fronting type is ruled out for vso-languages. It can only occur in languages whose basic word order type is either svo or SOV. Greenberg's (1966:83) Universal 11 states that inversion with polar interrogatives presupposes inversion with constituent interrogatives.

3.1.6 Verbal inflection Totally different from the strategies discussed so far is the way some polysynthetic languages like West Greenlandic (Kalaallisut), and Eskimo languages (Inuit) in general, encode polar questions (Sadock (1984:190)). As the minimal pair from West Greenlandic in (50) shows, the interrogative is derived from the declarative by morphological alternation. Additional languages possessing an interrogative mood marked by inflection are Tariana (Arawakan, Brazil), languages of the Tucano family spoken in Brazil/Colombia, and Blackfoot (an Algonquian language).

(50) West Greenlandic
      a. neri-vutit
         eat-IND.2SG.PAST
         'You ate'
      b. neri-vit
         eat-INT.2SG.PAST
         'Did you eat?'

3.2 *Constituent interrogatives*

With the exception of interrogative tags and disjunctive-negative structures, all the strategies used for deriving polar interrogatives as discussed in the previous section can also be found with constituent interrogatives. However, their occurrence is less wide-spread and their use, in many cases, optional. Constituent interrogatives, so it seems, can be more readily recognized as questions than polar interrogatives. One notable exception to this general impression is the morphological marking of questions as found in West Greenlandic (Sadock (1984:199)), where the same marker appears obligatorily in both polar interrogatives and constituent interrogatives:

(51) West Greenlandic
      a. neri-va-Ø?
         eat-INT-3SG
         'Did he eat?'

    b. su-mik   neri-va-Ø
        what-INSTR eat-INT-3SG
        'What did he eat?'

However, with respect to the remaining strategies, such a conclusion seems indeed justified. For example, of the thirty-six languages in Ultan's (1978) sample for which intonation as a means for marking polar interrogatives is attested (mostly rising intonation or higher pitch), only twelve (or one third) also use the same or a similar intonational pattern for marking constituent interrogatives. The overall impression that the data give is that most languages either do not mark constituent interrogatives by intonation at all (33.3 per cent) or do so only optionally (33.3 per cent). No intonational marking of constituent interrogatives is reported from Fula, Japanese and Tagalog, whereas languages like Amharic, English and Turkish at least have optional marking. A similar problem is the occurrence of interrogative particles in constituent interrogatives. From a functional point of view, it appears superfluous to have such particles in constituent interrogatives. The interrogative word should unambiguously type the relevant sentences as interrogative. What we can observe empirically, however, is that in approximately 50 per cent of the languages interrogative particles are optionally added to constituent interrogatives. In the Japanese example in (52) below, the particle *ka* is not obligatory and may be added or simply left out:

(52)   Japanese
        Dare-ga   kimasu (ka)?
        who-NOM come   INT
        'Who is coming?'

Finnish, by contrast, does not allow the interrogative particle *ko/kö* in constituent interrogatives:

(53)   Finnish
        Kuka(*ko) tulee   huomenna?
        who   come.3SG tomorrow.ESS
        'Who is coming tomorrow?'

Attempts have also been made, although with limited success, to correlate the position of interrogative words with the position of interrogative particles in polar interrogatives. C. L. Baker (1970:207), based on Greenberg's (1966) data, hypothesizes that only languages which locate interrogative particles, provided they have such particles, in clause-initial position permit interrogative words in positions other than those of the constituents they replace, i.e. the position of such particles predicts whether a language has *wh*-movement or not (initial interrogative particle $\Rightarrow$ *wh*-movement). Contrasting with this view is Cheng

(1997:13ff.) who argues, mainly on the basis of Chinese, that all so-called '*in situ* languages' possess special particles to mark constituent interrogatives, although this marking may be covert, and all languages with such particles are *in situ*. Put in a nutshell, the distribution of interrogative particles in constituent interrogatives is clearly governed by certain constraints, but so far it has not been possible to identify them precisely. What seems to be relatively uncontroversial, by comparison, is the assumption that interrogative tags do not occur in constituent interrogatives.

Whether the reordering of constituents should be considered important for the identification of constituent interrogatives across languages depends to a large extent on the perspective taken. Restricting the scope to cases of subject–verb inversion leaves very few languages as plausible candidates manifesting this phenomenon (mainly the Germanic languages). If, by contrast, we examine the position of interrogative words relative to the position of the constituents they substitute for, a much greater range of languages has the property in question. For instance, it is very common for languages, quite independently of their genetic affiliation, to place interrogative words in sentence-initial position. For some languages it is obligatory to do this kind of reordering (English, German, Hebrew, Supyire, Yoruba, Zapotec), others just show a strong tendency to do so (Egyptian Arabic, Kannada, Korean, Palauan). An example of such a fronting language (Finnish) is shown in (54); the example from Swahili in (55) illustrates constituent interrogatives of optional fronting languages (Haiman (1985:245)).

(54)   Finnish
      a. Maija   ottaa   omenaa.
         Maija   take.3sɢ apple.ᴘᴀʀ
         'Maija is taking an apple.'

      b. Mitä   Maija   ottaa?
         what.ᴘᴀʀ Maija   take.3sɢ
         'What is Maija taking?'

(55)   Swahili
      a. A-li-fika   lini?
         3sɢ-ᴘᴀsᴛ-arrive when
         'When did s/he arrive?'

      b. kwa nini chakula ki-me-chelewa?
         why   food   3sɢ-ᴘᴇʀꜰ-late
         'Why is the food late?'

In Mandarin Chinese, and also in Indonesian, Japanese and Lezgian, interrogative words remain exactly in the position of the constituent which they replace (so-called '*in situ* languages'). Example (56) illustrates this point for Mandarin:

(56)   Mandarin Chinese
    a. Hufei măi-le   yī-běn-shū
       Hufei buy-ASP one-CL-book
       'Hufei bought a book.'

    b. Hufei măi-le   shénme?
       Hufei buy-ASP what
       'What did Hufei buy?'   Cheng (1997:5)

The position of interrogative words depends, to a certain extent at least, on the basic word order type of a language. Greenberg (1966:82) found a systematic correlation between VSO order and fronted interrogative words as well as between SOV order and the *in situ* parameter. However, the correlation is much weaker in the case of SOV languages. No such correlation can be established for SVO languages. Another point that is interesting from a cross-linguistic perspective is the behaviour of languages when it comes to the co-occurrence of multiple interrogative words. Fronting languages behave surprisingly differently with respect to this parameter: some of them neatly stack interrogative words at the beginning of a sentence while others only front one interrogative word and leave the rest in the positions where they logically belong (i.e. *in situ*). As (57) and (58) demonstrate, English fronts only one interrogative word leaving additional ones *in situ*, whereas Russian, and Slavic languages in general, assemble them at the beginning of a sentence:

(57)   a. John gave the book to Mary.
    b. What did who give to whom?

(58)   Russian
    Kto kogo   ljubit?
    who whom   loves
    'Who loves whom?'

    Languages also show differences in their inventory of interrogative words. One usually finds interrogative words which replace the core constituents or arguments of a sentence and typically inquire about persons and things (*who* versus *what*) as well as interrogative words in an adverbial function which are typically used to seek information about (i) the location of a situation (*where*), (ii) its temporal setting (*when*), (iii) the manner of carrying it through (*how*), and (iv) the reason for it (*why*). Besides this core inventory one can also find interrogative words for determiners and/or adjectives (English *which*, Finnish *kumpi* 'which one of the two'), quantifiers (French *combien* 'how many') and ordinal numbers (Finnish *monesko*, German *der wievielte* 'the how many-th'). Specific interrogative words for verbs or verb phrases like Tahitian *eaha*, Tagalog *ano*,

Palauan *mekera* or Futunian *ā'* are frequently found in Oceanic languages but are rare otherwise. Interrogative words for prepositions have so far not been attested.

4   Imperative sentences

Even a superficial glance at the strategies of imperative formation encountered across the world's languages makes it clear that variation in this area is as least as extensive as in the case of interrogatives. It is certainly no exaggeration to say that most, if not all, languages have at least one strategy for identifying imperatives, i.e. constructions dedicated to the expression of directive speech acts, i.e. orders and requests, but also invitations, the giving of advice, warnings, wishes, instructions, etc. (W. D. Davies (1986:30ff.)):

(59)   a. Please clean the bath after use.
         b. Have some more cake.
         c. Take plenty of exercise if you want to slim.
         d. Watch out for the dog.
         e. Sleep well.
         f. Ring bell for service.

In most studies, the label 'imperative' is reserved for sentence types expressing such speech acts when they are directed to addressees in the narrow sense of the word (second person).[15] Some authors extend this narrow definition to include commands, requests, etc., addressed to the first and sometimes even to the third person (Xrakovskij (2001)), i.e. to cases for which, traditionally, labels like 'hortatives', 'optatives', 'jussives' and the like would be used. In keeping with the traditional definition, we understand imperatives as sentences with an understood second person subject.

The most widespread strategy for marking imperatives seems to be a special inflectional form of the verb so that the traditional Western approach of subsuming the imperative under the category 'mood' appears justified even from a cross-linguistic perspective (van der Auwera and Lejeune (2005a, 2005b); van der Auwera, Dobrushina, and Goussev (2004)). This includes cases of genuine imperative affixes, but also the use of the bare verb stem and special verb stems, as well as morphological marking taken over from different domains, as, for example, subjunctive, aorist (perfect) and passive forms. A fairly general characteristic of inflecting languages is that morphological marking

---

[15] A prototypical imperative implies a second person addressee. If there are further person distinctions made in the same paradigm they follow the following hierarchy (Aikhenvald (2003)): 2 > 1PL (inclusive) > 3SG/PL > 1SG and/or 1PL (exclusive).

of the imperative is less extensive in comparison to other moods (indicative, subjunctive, etc.). Agreement affixes for person, number and gender, but also those indicating tense, aspect and the like, are frequently suppressed. However, the picture is complicated by the fact that the kind of imperative strategy employed frequently depends on the polarity of the sentence. Many languages use one strategy in affirmative sentences, but resort to a totally different strategy in negative sentences (so-called 'prohibitives' or 'vetatives'; van der Auwera and Lejeune (2005b)). What is also extremely common, if not universal, is the suppression of the subject (pronoun) in imperatives. Although many languages allow the optional use of a subject pronoun, mainly for contrastive purposes or to soften the force of the imperative (politeness / degree of formality), its absence clearly represents the unmarked case. These general patterns notwithstanding, the range of variation found in the domain of imperatives is quite extensive. We will give an overview in the following paragraphs, starting with what appear to be the most widespread strategies for positive and negative contexts. This will be followed by some brief remarks about more indirect means of expressing directive force as well as related constructions.

4.1   *Positive imperatives*

One of the languages in our sample that possesses a true imperative marker, i.e. an affix exclusively dedicated to the expression of directive force, is Limbu, a Tibeto-Burman language spoken in Nepal (Driem (1987:188)). In Limbu the imperative marker ɛʔ is used in the singular and plural, and always occurs in word-final position:

(60)   Limbu
　　　　a. ips-ɛʔ
　　　　　　sleep-IMPER
　　　　　　'Sleep!'
　　　　b. ips-ɛtch-ɛʔ
　　　　　　sleep-2D-IMPER
　　　　　　'Sleep (you two)!'
　　　　c. ips-amm-ɛʔ
　　　　　　sleep-2PL-IMPER
　　　　　　'Sleep (all of you)!'

In many other languages, however, we find imperative markers of a fusional type which provide further information in addition to illocutionary type, usually person and number. The imperative suffixes -*a* (SG) and -*wch* (PL) of Welsh are used to issue directive speech acts to a single addressee or to a group,

respectively. Another such language is Acholi, a Nilo-Saharan language (Crazzolara (1955:110)):

(61)   Acholi
         a. lok-i
            turn-2SG.IMPER
            'Turn!'
         b. lok-wu
            turn-2PL.IMPER
            'Turn (pl.)!'

In Finnish, by contrast, the imperative marker *kaa/kää* only appears in the plural (62). In the singular, the imperative in Finnish has the same form as the first person singular of the present indicative without the final *–n* (63).

(62)   Finnish
         sano-kaa
         talk-2PL.IMPER
         'Talk (pl.)!'

(63)   Finnish
         a. sano-a
            talk-INF
         b. sano-n
            talk-1SG.PRES.ACT
            'I talk.'
         c. sano
            talk.2SG.IMPER
            'Talk (sg.)!'

The imperative markers of Evenki, a Tungus language spoken in Siberia, also encode person and number (Nedjalkov (1997:18ff.)):

(64)   Evenki
         Purta-va-s   min-du bu:-kel
         knife-ACC.DEF-2SG.POSS   I-DAT   give-2SG.IMP
         'Give (you.sg.) me your knife!'

The relevant Evenki forms for all combinations of person and number are shown in Table 5.3. What should be borne in mind, however, is that the forms of the first and third person perform a somewhat different function. Morphological markers expressing commands or requests to the first person are usually called 'hortatives', those addressing third persons are sometimes referred to as 'optatives' or 'jussives'. In addition to the paradigm shown in Table 5.3, Evenki has

Table 5.3 *The imperative paradigm of Evenki*

1SG	*baka-kta*	'let me find'
2SG	*baka-kal*	'find'
3SG	*baka-gin*	'let him/her find'
1PL.EXCL	*baka-kta-vun / baka vvun*	'let us find'
1PL.INCL	*baka-gat*	'let us find'
2PL	*baka-kallu*	'(you pl.) find'
3PL	*baka-ktyn*	'let them find'

another complete paradigm for the encoding of orders, requests, commands, etc., that need not be executed immediately. Thus, what we find grammaticalized in Evenki is the distinction between a near future and a remote future imperative; cf. (65) vs. (66). At the same time, and quite plausibly, the remote forms are used as the polite imperative.

(65)   Evenki
        D'u-la-vi   himat   eme-kel
        home-ALL-POSS quick(ly) come-2SG.IMPER
        'Come quickly to my place!'

(66)   Evenki
        D'u-la-vi   (gochin)   eme-de:-vi
        home-ALL-POSS (next.year) come-IMPER-POSS
        'Come to my place (next year)!'

Quite unexpected, at least from a European perspective, is the case of Macushi, an Amazonian language of the Carib family (Abbott (1991:49ff.)), where there are special imperative markers indicating motion toward or away from the speaker, note the contrast between (67) and (68). A movational imperative is also reported from Chontal (Hokan).

(67)   Macushi
        apo' era'ma-ta,   ta-'pî-i-ya
        fire go-IMPER.MOT say-PAST-3-ERG
        '"Go get firewood," he said.'

(68)   Macushi
        tuna   era'ma-tane'kî, ta-'pî   i-san-ya
        water get-IMPER.MOT say-PAST 3-mother-ERG
        '"Come get the water," her mother said.'

What appears also quite common across the world's languages is to have no special imperative marker at all and simply to use the bare verb stem. In many of these cases, however, a plural marker appears with commands directed to a group of addressees. This asymmetry corresponds to well-established markedness patterns according to which the singular represents the unmarked category (Croft (1991)). Imperative formation of this kind is found in Turkish (Kornfilt (1997:41)), German, Persian, Punjabi, Shona and many other languages. The relevant paradigms for Turkish and German are given in (69).

(69)   Turkish   German
    a. git-mek   komm-en
        go-INFIN   come-INFIN

    b. git   komm
        go.IMPER.2SG come.IMPER.2SG

    c. gid-in   komm-t
        go.IMPER-2PL come.IMPER-2PL

In Spanish the imperative verb form of the second person singular is identical to the verb form of the third person singular present indicative. In the plural the suffix *-d* is added, which yields an unambiguous imperative form:

(70)   Spanish
    a. canta
        sing.2SG.IMPER (= 3SG.PRES.INDIC)

    b. cantad
        sing-2PL.IMPER

Gulf Arabic is one of the few languages in our sample that draws a gender distinction in imperatives. This, however, is done in the singular only, again in complete harmony with markedness patterns (71). Another such language is Abkhaz.

(71)   Gulf Arabic
    a. ʔiktib-Ø
        write-MASC

    b. ʔiktib-i
        write-FEM

    c. ʔiktib-u
        write-PL

Special imperative stem forms can be found, *inter alia*, in German, Lezgian and Welsh, but whereas the relevant German stems are clearly related forms (see

(72), we find true suppletion in Lezgian see (73). A few imperatives of Lezgian are derived from the stem by consonant reduplication.

(72)   German
        a. nehmen(INFIN   nimm(IMPER)
           'take'   'Take!'

        b. geben(INFIN)   gib(IMPER)
           'give'   'Give!'

(73)   Lezgian
        a. atu-n(INFIN)   ša(IMPER)
           'come'   'Come!'

        b. fi-n(INFIN)   alad(IMPER)
           'go'   'Go!'

Finally, it should be mentioned that many languages do not possess morphological markers dedicated to expressing imperative force. An obvious example is English, together with many languages of Southeast Asia (van der Auwera and Lejeune (2005a)).

4.2   *Negative imperatives (prohibitives)*

Negative imperatives deserve special mention since they may be similar to, but also very different from, their positive counterparts. There are basically four strategies according to which languages encode negative directive speech acts (van der Auwera and Lejeune (in press b)):

(i) the use of a positive imperative verb in combination with the negative strategy found in declaratives;

(ii) the use of a positive imperative verb in combination with a negative strategy not found in declaratives;

(iii) the use of a verb other than the positive imperative and the negative strategy found in declaratives;

(iv) the use of a verb other than the positive imperative and a negative strategy not found in declaratives.

In addition, there is some minor variation on top of these basic distinctions. Some languages even have special prohibitive morphology, i.e. affixes expressing negative directive speech acts without the relevant sentences being overtly negative.

The first major strategy introduced above can be illustrated with Turkish and German, where negative commands are expressed by using the verb stem in combination with the negative marker found in declarative sentences (74). As pointed out above, only the bare verb stem is used for positive imperatives addressed to one person (cf. (69)).

(74)   Turkish   German
        git-me   komm   nicht
        go-NEG   come   NEG
        'Don't go (sg.)!' 'Don't come (sg.)!'

Somewhat different, but still comparable to cases like Turkish and German, are languages that require a special auxiliary in negative contexts. The *do*-periphrasis known from English would be a case in point. What is striking about English is that negative imperatives require *do*-support even with the verb *be*, which can be combined with a following negation without *do*-support in declaratives and interrogatives:

(75)   Don't be a fool.

Evenki is quite similar to English in this respect (also the Carib language Wai Wai), but adds the imperative suffix to the negative auxiliary *e-* 'not to' (76). It should be stressed that the strategies of negation illustrated in (75) and (76) represent the normal means of sentence negation and have nothing to do with imperatives proper.

(76)   Evenki
        Tala   e-kel   girku-ra
        there NEG.AUX-2SG.IMPER   go-PART
        'Don't go there!'

As for the second major strategy, i.e. the use of a special negative marker together with the positive imperative verb, this can be illustrated with the Vietnamese data in (77), where in negative imperatives the negative element *chó* appears, in contrast to the normal sentence negation *không* (cf. L. C. Thompson (1965:210, 221)). Another such language is Punjabi (Bhatia (1993:40)).

(77)   Vietnamese
        a. Chó uông   ruou!
            NEG drink   alcoholic.beverages
            'Do not drink alcoholic beverages!'

        b. Tôi không hieu
            I   NEG   understand
            'I do not understand.'

A related, but slightly different strategy can be found in Welsh, Finnish, and Samoan, where imperatives are negated by means of special negative auxiliaries which only occur in imperatives (in contrast to English and Evenki above). Example (78) is from Finnish, (79) from Samoan (Mosel and Hovdhaugen (1992:482)). Note that Finnish marks both the main verb and the (negative)

auxiliary as imperative (but only in the plural). The special imperative marker *-ko* is formally indistinguishable from the interrogative particle.

(78)   Finnish
    Äl-kää   tul-ko
    NEG.AUX-2PL.IMPER   come-IMPER
    'Don't come!'

(79)   Samoan
    'Aua e   te   fa'asāunoa 'i mea-ola
    don't 2SG ASP torture   to thing-life.PL
    'Don't torture animals!'

An example of a language with the third major strategy introduced above (negation of declaratives plus verb form different from the positive imperative) is Spanish, where negative commands, requests, etc. are expressed with the relevant verb forms in the subjunctive in combination with the normal sentence negation:

(80)   Spanish
    a. canta   cantad
       sing.2SG.IMPER sing.2PL.IMPER
       'Sing!'   'Sing (all of you)!'
    b. no   cantes   no   cantéis
       NEG sing.2SG.PRES.SJNCT NEG sing.2PL.PRES.SJNCT
       'Don't sing!'   'Don't (you all) sing!'

In Italian we find the infinitive used for negative imperatives, albeit only in the singular:

(81)   Italian
    a. canta   cantate
       sing.2SG.IMPER sing.2PL.IMPER
       'Sing!'   'Sing (all of you)!'
    b. non cantare   non cantate
       NEG sing.INFIN NEG sing.2PL.IMPER
       'Don't sing!'   'Don't (you all) sing!'

Finally, in the fourth major strategy of forming negative imperatives, a form of the verb other than the one found in positive imperatives is used together with a negative element which does not occur in declarative sentences. Examples of this strategy are found in Malagasy (see Dez (1980:33, 167)), where negative imperatives are formed with the negative element *aza* (the declarative negation is *tsy-*) and a verb in the indicative (82), as well as in Kannada (Sridhar (1990:36)),

which in negative imperatives requires a verb in the infinitive and a negative auxiliary instead of the normal sentence negators *alla* or *illa* (83):

(82)   Malagasy
      a. Tongav-a!
         come-IMPER
         'Come!'

      b. Aza mitomany!
         NEG cry.INDIC
         'Don't cry!'

(83)   Kannada
      A: ka:De ho:g-a-be:Da!
      that side   go-INFIN-NEG
      'Don't go that way!'

One of the few languages in our sample to possess an unequivocal prohibitive marker is Lezgian (Haspelmath (1993:149f.)). In this language (as well as other Caucasian languages), the imperative marker *-a* stands in paradigmatic opposition to the prohibitive marker *-mir* – cf. the contrast between (84a) and (84b). Such prohibitive sentences do not contain sentential negation.

(84)   Lezgian
      a. Wuna   bağišlamiš-a, buba
         you.ERG forgive-IMPER father
         'Forgive me, father!'

      b. Wa-z   kič'e   že-mir
         you.DAT afraid   be-PROHIB
         'Don't be afraid!'

Although true prohibitive markers appear to be a comparatively infrequent phenomenon – there are just four languages in our sample (of about seventy languages, see note 3) that have a special morphological marker for the expression of negative directive speech acts (Lezgian, Macushi, Malayalam, Warekena) – the previous discussion has shown that it is relatively common for languages to treat negative imperatives differently from positive imperatives in one way or another. Overall, imperatives tend to preserve archaic forms and exhibit less compositionality than the other sentence types.

4.3   Indirect strategies

For a few languages in our sample it has not been possible to identify a construction uniquely dedicated to the expression of directive force. What we also found are languages that in principle have a true imperative construction, but

only very rarely make use of it and are gradually replacing it with a construction taken from a different domain. Such is the case in Modern Hebrew, where verb forms marked for future tense are the normal or unmarked means of expressing directive speech acts (85), even though a true morphological strategy is available. The use of the future tense is obligatory for negative commands, since the imperative is not possible in such contexts (Glinert (1989:284ff.)). Another language well known for avoiding its genuine imperative strategy is English (cf. *Thank you for not smoking*, *Could you . . .?*).

(85)   Modern Hebrew
         Te-   sader
         FUT-   tidy
         'Tidy!'

That languages use markers from the domain of tense and aspect for the expression of directive force is also attested in other languages. For example, the imperative of Georgian is formally indistinguishable from the aorist, and in Rapanui, a Polynesian language (Du Feu (1996:37ff.)), directive force is expressed by normal declarative sentences in the present tense combined with a temporal adverb meaning something like 'now' or 'just', i.e. a momentary temporal unit (86). The literal translation into English makes clear that such sentences can indeed be understood as commands: 'Now you wipe your face'. Obviously, the distinction between direct and indirect speech acts is not applicable to languages like Rapanui.

(86)   Rapanui
         Ka   amo   te   'arinŋa
         now   clean DET   face
         'Wipe your face!'

Similar observations can also be made for German, where the perfect participle can be used to express commands of a rather impolite kind: *Jetzt aber aufgestanden!* 'Get up!' (lit.: 'now but got up'). As a matter of fact, it is quite conceivable that temporal and aspectual markers are a source, maybe even a major one, for the grammaticalization of imperatives.

Another indirect strategy of imperative formation is the use of subjunctives. The imperative of Lango, a Nilo-Saharan language spoken in Uganda (Noonan (1992)), is formed by dropping the subject agreement affixes from the relevant subjunctive forms. And in Hungarian (cf. Kenesei, Vago, and Fenyvesi (1998)), as well as in Persian, the imperative marker (-*j* and *be-*/*bo-* respectively) is formally equivalent to the subjunctive marker; cf. the Hungarian examples in (87a) and (87b).

(87)   Hungarian
       a. Másol-j   egy kulcs-ot
          copy-IMPER a   key-ACC
          'Copy (2sg) a key!'

b. Nem szükséges, hogy Péter meg-tanul-j-a   a   vers-et
   NEG necessary that   Peter COMPLET-learn-SJNCT-DEF.3SG the poem-ACC
   'It isn't necessary for Peter to learn the poem.'

The exact status of the imperative marker in Hungarian is a matter of some
debate, but given the core meaning of the subjunctive, it is certainly not unex-
pected to find its range of uses extended in this way. By comparison, it appears
more challenging to explain why the passive can also be used as an indirect
strategy for imperatives, as is the case in Maori, a Polynesian language (Bauer
(1993:32)).

(88)   Maori
       Patu-a   te   kurii raa
       beat-PASS the dog   DIST
       'Beat that dog!'

4.4   Related constructions

Closely related to imperatives, i.e. constructions expressing directive speech
acts such as commands, requests, advice, suggestions, invitations, etc., are for-
mal markers frequently referred to as 'hortatives', 'optatives', 'debitives', 'rog-
atives' and 'monitories', which are typically associated with the illocutionary
forces given in Table 5.4. Moreover, there is a difference in person associated
with some of these labels: the label 'imperative' is often restricted to second
person directives, whereas 'hortative' is found for first and third person direc-
tives and 'optative' for directions addressed to third persons. In the strict sense,
which is not usually applied, a language is said to possess any of these categories
if it has special verbal morphology exclusively dedicated to the expression of
the relevant illocutionary functions. Such inflectional markers do indeed exist
in many languages other than English. In principle, one could imagine a lan-
guage with full morphological paradigms for each of these functions, maybe
even inflecting for person, number, etc. In practice, this happens very rarely, if
at all. At least, we are aware of no such language. What does happen is that
languages pick a seemingly idiosyncratic mix of forms with often overlapping
functions. An additional complicating factor is that the usage of these terms is
not completely fixed and varies from author to author.[16]

---

[16] In particular, the frequent confusion of form and function makes information provided in gram-
matical descriptions difficult to interpret.

Table 5.4 *Subcategories of imperatives*

Category	Illocutionary force
hortatives	exhortations
optatives	wishes
debitives	obligations
rogatives	petitions
monitories	warnings

Of course, apart from morphological marking, there are always alternative means of expressing the relevant illocutions. The English construction for the expression of exhortations, i.e. commands to the first person or a group of people including the speaker, is *let's*, as in *Let's go to the movies*. Wishes, obligations and warnings are either expressed by explicit performatives (*I wish . . .* ; *I warn . . .*) or by modal verbs (*You must leave now*). True morphological hortatives are, *inter alia*, found in Lezgian (89), Evenki (90) and French (91).[17]

(89)   Lezgian
       Sifte wun   wi   buba.di-z   ̄qalur-in
       first you.ABS you.GEN father-DAT show-HORT (=1SG.IMPER)
       'First let me show you to your father!'

(90)   Evenki
       Bi oro-r-vi   baka-kta
       I reindeer-PL-POSS find-HORT (= 1SG.IMPER)
       'Let me find my reindeer.' / 'I'll go and find my reindeer.'

(91)   French
       a. chant
          sing.2SG.IMPER
          'sing!'

       b. chantez
          sing.2PL.IMPER
          'sing!'

       c. chantons
          sing.1PL.IMPER (= HORT)
          'Let's sing.'

Examples of languages with optatives, i.e. morphological markers expressing wishes (or third person directives), are Malayalam, Lezgian, Evenki, Greek,

---

[17] Strictly speaking, the form *chantons* in (91c) is first person plural present indicative (*nous chantons*). What makes it imperative or hortative is the omission of the subject.

Turkish. An example from Mayalayam is shown in (92), and from Lezgian in (93).

(92)   Mayalayam
       avar samsaarikk-aṭṭe
       they speak-OPT
       'Let them speak.'

(93)   Lezgian
       Quj wun   či   Cükwer.a-z   wax   x̂u-raj
       let you.ABS we.GEN   Cükwer-DAT   sister be-OPT
       'May you be a sister for our Cükwer.'

Among the languages in our sample possessing debitives, i.e. verbal morphology reserved for the expression of obligation, are Malayalam (94) and Evenki (95). The relevant markers -*aŋam* and -*mechin* inflect neither for person nor for number.

(94)   Mayalayam
       niŋŋaḷ naaḷe   taṉṉe var-aṇam
       you   tomorrow EMPH   come-DEBIT
       'You must come precisely tomorrow'

(95)   Evenki
       Minngi girki-v   ilan-duli   chas-tuli   suru-mechin-in
       my   friend-1SG.POSS three.PROL hour.PROL   go.away-DEBIT-3SG
       'My friend must leave in three hours.'

The decision whether a language has a rogative (for petitions) or not depends to a large extent on the definition of this category. If it is defined broadly so as to comprise polite imperatives, then a not insubstantial number of languages will qualify for inclusion. We have already mentioned that Evenki has two complete imperative paradigms, one expressing categorical/immediate commands, the other one being reserved for those that are polite/remote. Similarly, the polite imperative -*watá* of Chontal (Hokan) could plausibly be called a rogative:

(96)   Chontal
       ɬnáy-watá
       'Please, let me know.'

The final subcategory of imperatives to be discussed in this section is the category of 'monitory', i.e. verbal inflection used for the expression of warnings. One of the few languages possessing a true monitory is again Evenki (97). Of course, there are numerous indirect strategies for expressing the same illocutionary function. For instance, warnings in Malayalam can be expressed by combining the imperative marker with a negative tag.

(97)   Evenki
         Er-tyki,   tar-tyki   iche-t-ne
         this-ALL   that-ALL   see-CONTIN-MONIT
         '(Be careful and) look in different directions.'

It is an intriguing question whether there are any systematic relations between the subcategories of the imperative and also between these subcategories and the imperative itself. A plausible assumption is that the distribution of prohibitives, hortatives, optatives, etc., is not totally random but subject to implicational generalizations of the kind that if a language has category X, it will also have category Y. In the ideal case, the categories under consideration here would permit a ranking on an implicational hierarchy. Unfortunately, only a minority of languages in our sample draw the relevant morphological distinctions so that the formulation of any such strong hypothesis cannot even be attempted. Nevertheless, what our data clearly show is that some of these categories are very likely to occur together – although it is not clear in which order – and that there are also some implicational connections. For instance, Macushi and Malayalam have four of these categories; Lezgian, Turkish, Evenki and Wai Wai have three; and Georgian has special morphological markers for at least two of them. Moreover, the existence of any of these subcategories implies the existence of a true (morphological) imperative in a language (cf. van der Auwera *et al* (2004)).

5   Some minor sentence types

5.1   *Exclamatives*

Among the minor sentence types that can be distinguished across languages, in addition to the three major ones and their subtypes, exclamatives are the most prominent.

In terms of the typology of speech acts mentioned above, exclamations, the semantic counterpart of so-called 'exclamative sentences', are used for the performance of representative speech acts, i.e. for speech acts expressing a state of belief and making a claim about the world. But in contrast to assertions, the point of an exclamation is not really to inform the hearer(s) about some situation, but to express an affective response to what is taken to be a fact. More specifically, exclamations convey the speaker's surprise that some present situation is remarkable and thus seem to be used as expressive speech acts of a type not included in Searle's typology. Finally, exclamations relate to a scale or dimension and identify an extreme value.

Exclamations can be expressed by a wide variety of formal structures and constructions (cf. Rosengren (1992b); Michaelis (2001)). In English, declarative

sentences, interrogative sentences, free relatives, isolated NPs, inversion, subordination to factive epistemic verbs may be used, *inter alia*, for that purpose:[18]

(98) a. He is so stupid / such an idiot!
      b. Isn't this great!
      c. How foolish he is!
      d. The speed they drive on the freeway!
      e. Man, is this kid intelligent!
      f. I can't believe how much he has grown!

A similar variety of structures can be found in many other languages (cf. Michaelis (2001)). In addition to the structures mentioned for English, declarative sentences introduced by the complementizer *dass* 'that' can be used in German, as well as all free relatives introduced by an interrogative pronoun:

(99) German
      a. Dass der immer nur   Tennissocken trägt!
          that he   always only tennis.socks   wears
          'It is incredible that he always wears tennis socks.'

      b. Wen   die   alles eingeladen haben!
          whom they all   invited   have
          'The people they invited!'

Given this variety of structures and constructions that can be used to express exclamations, none of which can easily be dismissed as being an example of an indirect speech act, it is, of course, highly problematic to list exclamatives as a fourth basic sentence type alongside declaratives, interrogatives and imperatives. The only common denominator of all these structures seems to be intonation: exclamative constructions are generally characterized by a falling intonation contour and a focus on either the basic argument or the (scalar) predicate or on both. In view of these facts it seems justified to exclude exclamatives from the list of basic sentence types (cf. Rosengren (1992b:265f.)). So-called 'exclamative sentences' could simply be regarded as being the result of combining declarative or interrogative sentences with specific syntactic, semantic and pragmatic properties, all of which are highly suitable and thus motivated for the expression of an exclamation. Zanuttini and Portner (2003) thus restrict the term 'exclamative' to those sentences that contain an interrogative word and are factive. A different view is provided in one of the very few typological studies of exclamative constructions (Michaelis (2001)). Michaelis regards exclamatives as constructions, i.e. as a set of semantico-pragmatic features, all of which must receive formal expression.

---

[18] Note that there are two properties that distinguish (98b) from (98a): (98b) invariably contains a negation and can have rising intonation; structures of type (98a) have neither of those properties.

5.2   *Echo questions*

Somewhat different from their central use is the use of questions for the purpose of seeking clarification on a preceding utterance. Although both polar interrogatives and constituent interrogatives easily lend themselves to this purpose, echoing something previously uttered in the form of a question often precludes the use of the standard interrogative constructions and necessitates a more indirect way of asking questions. In English, for instance, polar echo questions show a strong tendency to be phrased in terms of a declarative sentence with a rising intonation. Moreover, using constituent questions as echoes means that interrogative words can and frequently are left *in situ*:

(100)   A: I tell you he is a braggart.
          B: He is a braggart? / He is what?

As a matter of fact, more careful analyses of echo questions have revealed that their discussion in the context of interrogative constructions and their frequent incorporation into this domain is problematic. One of the major problems for such an apparently straightforward approach is that all the three basic sentence types can be used as echo questions. The examples given in (101) show declarative, interrogative and imperative sentences being used as echo questions. The only property that these echo questions have in common with interrogative sentences is that they contain an interrogative word (constituent echo questions at least). These and similar facts have led some authors to conclude that echo questions cannot be regarded as instantiating a special sentence type, but are a phenomenon that can be superimposed on any of the basic sentence types (cf. the discussion in Reis (1992); Huddleston (1994)).

(101)   a. A: John lives in Paris. – B: He lives where?
          b. A: John lives in Paris. – B: Where does he live?
          c. A: Go to Paris! – B: Go where?

Another point of some debate is whether echo questions are used to perform independent speech acts or whether they merely quote the preceding utterance. What, among other things, argues in favour of a quotational analysis is that echo questions do not represent the point of view of the speaker. The evaluation of a certain person's mental capacity expressed by B in (102) is evidently not that of the speaker:

(102)   A: I have met this idiot again.
          B: You have met which idiot again?

Moreover, such an analysis is corroborated by languages which regularly use a quotative construction to express echo questions, as, for example, Turkish (103). The quotative verb is not optional in this example (cf. Kornfilt (1997:32)).

(103)   Turkish
      A: Sinema-ya   gid-iyor-um
         cinema-DAT go-PRES.PROG-1SG
         'I am going to the movies.'

      B: Sinema-ya   gid-iyor-um   mu de-di-n
         where-DAT   go-PRES.PROG-1SG INT say-PAST-2SG
         'Did you say "I am going to the movies"?'

In all the languages we surveyed, echo questions may be reduced, and in fact
are often reduced, to the constituent or the constituents on which clarification
is required. In informal speech it is probably even more common simply to
use the interrogative word for inanimate referents (English *what*). Somewhat
different from the examples of echo questions just discussed are those that are
used as a reply to a preceding question. In German and English, at least, such
echo questions have the structure of embedded clauses and cannot be analysed
as quotations. Consider the example of a polar echo question in (104B) and the
one of a constituent echo question in (105B).

(104)   German
      A: Hast du   schon   eingekauft?
         have you already shopped
         'Have you done the shopping yet?'

      B: (Du fragst,) ob ich schon   eingekauft habe?
         you ask   if I   already shopped   have
         '(You are asking) if I have done the shopping?'

(105)   German
      A: Wo   bist du   gewesen?
         where are you been
         'Where have you been?'

      B: (Du fragst,) wo   ich gewesen bin?
         you ask   where I   been   am
         '(You are asking) where I have been?'

5.3   *Nonfinite presentatives*

Another minor sentence type that can be identified across languages are non-
finite constructions that are not declarative, interrogative, or imperative and that
are used to present a proposition with the purpose of rejecting it as absurd:

(106)   a. English
         Him play the piano. Ludicrous!

         b. German
         Der und Klavier spielen. Lächerlich!
         this and piano   play   Ludicrous

Constructions like (106) have been discussed under such labels as 'MAD maga-zine sentences', 'left dislocations of argument and predicate', 'sentential topics', 'citations', etc. The most interesting point about such constructions is that they are not tokens of one of the three major sentence types and could be analysed as instantiating a 'presentative mode' (cf. Zaefferer (1990:223f.)). It is probably due to the marginal role of such constructions in communication that they have not been accorded that status.

5.4   Answers to questions

Among the numerous sentence fragments that are used across languages, answers are easy to identify and also lend themselves to some noteworthy cross-linguistic generalizations. Answers are declarative sentences with a spe-cific focus marking. More often than not they are relatively short, reduced to their focus, and normally do not give more than the requested information, i.e. a truth value in the case of polar questions and the information specified by the interrogative word(s) in the case of constituent questions (107). Of course, more elaborate responses can also be encountered.

(107)   a. A: Are you leaving. – B: Yes (I am).

         b. A: Where are you going? – B: (I am going) to Paris. /
         Well, what do you think?

There seem to be three different answering systems for polar interrogatives: (i) yes/no systems, (ii) agree/disagree systems, and (iii) echo systems, the main properties of which can be described as follows. In yes/no systems of the type employed in English, German, Turkish, etc., confirmation of the proposition expressed by the relevant question is indicated by supplying an answer of the same polarity, whereas the polarity of the answer is opposite to that of the question in the case of non-confirmation. This applies to both positive and negative questions.

(108)   a. A: Did he bring a present? – B: Yes. (confirmation)
         b. A: Did he bring a present? – B: No. (non-confirmation)

(109)
a. A: Did he not bring a present? – B: Yes, he did. (non-confirmation)
b. A: Did he not bring a present? – B: No. (confirmation)

There is no difference between yes/no systems and agree/disagree systems as far as questions of positive polarity are concerned. Where the two systems diverge is in negative contexts, the confirmation/non-confirmation pattern of the respective answers being exactly reversed. The example in (110) simulates an agree/disagree system on the basis of English.

(110)
a. A: He did not bring a present, right? – B: Right. (confirmation)
b. A: He did not bring a present, right? – B: Wrong. (non-confirmation)

Languages possessing agree/disagree systems in our sample include Gulf Arabic, Japanese, Malayalam and Punjabi, among others. For instance, in Gulf Arabic *naʕam* is the agreement particle whereas *bala* is used to express disagreement:

(111)   Gulf Arabic
        A: maa   ʕindik   fluus,   muu chidhi
           NEG   with.you money NEG like.that
           'You haven't any money, right?'

        B: naʕam 'It is true that I have no money.'
           bala   'It is not true that I have no money' i.e. 'I have money.'

No special answer words at all can be found in the third type of answering system, i.e. the echo system, which works by using part of the question – usually the verb – as the answer. Welsh and Finnish are among the language in our sample possessing such an echo system (see the Welsh example in (112)). A special negative element (*na(c)*) is used in cases where the question cannot be answered affirmatively.

(112)   Welsh
        A: A   welwch chwi hwy?
           INT see   you   them
           'Do you see them?'

        B: Gwelaf   '(Yes) I see (them).'
           see
           Na   welaf '(No) I don't see (them).'
           NEG   see

One well-known problem of yes/no systems is that positive answers to negative questions can be confusing whenever the answer expected is biased toward

an affirmation. As the example in (113) illustrates, answering such questions simply by 'yes' leaves open whether what was meant is 'Yes, he did' or 'Yes, he didn't'.

(113)   A: He didn't bring a present, did he? – B: Yes.

Evidently, this ambiguity arises because *yes* is mistakenly interpreted as a marker indicating confirmation, i.e. as part of an agree/disagree system. In order to make it unambiguously clear that a positive answer to a negative question is intended and that the expectations raised by the question are wrong, many languages with yes/no systems offer a third answering strategy besides 'yes' and 'no'. For German this is *doch* (as opposed to *ja*), for French *si* (instead of *oui*), for Tigrinya *Ɂəbba* (rather than *Ɂəwwa*), and in English one can use a tag answer (e.g. *Yes, he did*) to achieve the desired effect.

6   Summary and conclusion

The preceding discussion has shown that the communicative potential of a sentence, i.e. the potential for performing actions (speech acts) of various kinds, is consistently and pervasively encoded in the grammar of languages, even if not typically in terms of paradigmatic oppositions. In particular the distinction between three basic sentence types, declarative, interrogative and imperative, is overtly drawn in most, even if not in all, languages. So-called 'exclamative sentences', by contrast, do not seem to constitute a separate basic sentence type, but can simply be analysed as the result of combining declarative or interrogative sentences with specific syntactic, semantic and pragmatic properties. Cross-linguistically, the three major sentence types are characterized by a limited set of recurrent strategies. Among these the imperative exhibits the highest degree of further differentiation, a fact which could find an explanation in the interactional risks associated with directive speech acts. Formal differentiation between imperatives and prohibitives, hortatives, optatives, rogatives, debitives, etc., is a fairly wide-spread phenomenon among the languages of the world. As far as interrogative sentences are concerned, most languages seem to distinguish polar interrogatives from constituent (*wh-*) interrogatives and also the unmarked use of both from the echoic use ('echo questions'). Declarative sentences are often identified by formal markers that conflate indicative or declarative mood with modal notions, such as evidentiality and strength of assertion. Note, however, that these basic sentence types are compatible with a wide variety of specific uses or speech acts. So, strictly speaking, what we find in the grammar of a language are general distinctions of sentence types, semantic mood or illocutionary potential, rather than 'speech act distinctions'.

Our findings, as those of others before us, show that the three major sentence types traditionally distinguished for European languages can also be clearly identified in a wide variety of other languages and that further differentiations are typically based on these primary form types. Our findings therefore suggest that the distinction between declarative, interrogative and imperative can and ought to play an important role in grammatical theory and in the analysis of the interface between grammar and pragmatics. What our survey also reveals is the fact that the semantic analysis of these categories must be a very abstract one. It is only as a result of the interaction of these basic sentence types with a variety of other formal, semantic and contextual properties that an utterance has a specific use or function in a context. Some of these interacting features, such as intonation for instance, seem to function very similarly across languages. Others seem to be language-specific (such as modal particles in German). And such combinations of basic sentence types with bundles of language-specific features may develop into constructions whose use potential can no longer be derived from the interaction of all features in a compositional fashion.

7   Suggestions for further reading

There are very few comprehensive cross-linguistic studies on the form and function of sentence types. Apart from the present chapter there is only its counterpart in the 1985 edition (Sadock and Zwicky (1985)) and Palmer's (very readable) volume on mood and modality (1986). Some methodological problems relating to the identification of sentence types are discussed in Croft (1994).

Moreover, there are a few studies on specific sentence types (as well as related phenomena) across different languages, although the number of studies available is surprisingly low. Ultan (1978) is the classic typological article on interrogatives. Chisholm, Milic, and Greppin (1984) contains useful descriptions of interrogative constructions from various languages (Japanese, Russian, Ute, etc.). Cheng (1997) offers a cross-linguistic study of *wh*-questions in the framework of government and binding. A recent summary of the cross-linguistic properties of interrogative constructions can be found in Siemund (2001). The main source for imperatives and related constructions (hortatives, optatives, etc.) are Xrakovskij (2001), van der Auwera and Lejeune (2005a, 2005b), as well as van der Auwera *et al*. (2004). A new major cross-linguistic study of imperatives is Aikhenvald (2003). Cross-linguistic aspects of exclamative constructions are discussed in Michaelis (2001). Declaratives are usually discussed in connection with markers of evidentiality: Chafe and Nichols (1986), Johanson and Utas (2000).

As for descriptions of sentence types in individual languages, the single best investigated language is probably German, with the two extensive volumes edited by Rosengren (1992a, 1993) still forming the major point of orientation. More recent discussions of German sentence types can be found in Reis (1999) and Lohnstein (2000). For English interrogatives and questions, useful starting points are Pope (1976) and Huddleston (1994). E. Davies (1986) provides detailed information on English imperatives. Comprehensive descriptions, however, are only available for relatively few languages (even among the well-studied languages) so that an approach via reference grammars is mostly inevitable.

Theoretical problems pertaining to the form–function relationship of sentence types are discussed *inter alia* in Sadock (1974), Searle, Kiefer, and Bierwisch (1980), Levinson (1983), Zaefferer (1990), Tsohatzidis (1994). The classic (still widely discussed) studies on speech acts and the function of utterances are Austin (1962) and Searle (1969) as well as Searle's subsequent publications.

6   Passive in the world's languages

Edward L. Keenan and Matthew S. Dryer

0   Introduction

In this chapter we shall examine the characteristic properties of a construction
wide-spread in the world's languages, the passive. In section 1 below we discuss
defining characteristics of passives, contrasting them with other foregrounding
and backgrounding constructions. In section 2 we present the common syntactic
and semantic properties of the most wide-spread types of passives, and in section
3 we consider passives which differ in one or more ways from these. In section 4,
we survey a variety of constructions that resemble passive constructions in one
way or another. In section 5, we briefly consider differences between languages
with regard to the roles passives play in their grammars. Specifically, we show
that passives are a more essential part of the grammars of some languages than
of others.

1   Passive as a foregrounding and backgrounding operation

Consider the following sentences:

(1)   a. Mary slapped John
        b. John was slapped
        c. John was slapped by Mary

Functionally speaking, passives such as (1b) and (1c) may be considered fore-
grounding constructions compared with the syntactically less marked and prag-
matically more neutral active, (1a): they 'topicalize' ('foreground', 'draw our
attention to') an element, *John*, which is not normally presented as topical
in the active. To this extent passives are similar to what we shall here call
topicalizations, (2b) below, and left-dislocations, (3b) below, both prominent
foregrounding constructions across the world's languages.

(2)   a. I like beans
        b. Beans I like

(3)   a. Congressmen don't respect the President any more
        b. As for the President, congressmen don't respect him any more

Functionally, the passives differ from these sentences in at least two ways. First, by eliminating the subject of the active, as in (1b), or by relegating it to the status of an oblique NP, as in (1c), they background the active subject in ways in which the topicalizations or left-dislocations do not.

Moreover, the passives seem to be weaker foregrounding constructions than either the topicalizations or the left-dislocations. Thus in (3b) *the President* is somehow more of a topic than is *congressmen*, the subject (= unmarked topic) of (3a). But in *John was slapped*, *John* seems to be a topic only to the same extent that *Mary* is in the corresponding active, *Mary slapped John*. Notice that it is generally quite difficult across languages to topicalize or left-dislocate twice from the same sentence (some exceptions are known). Thus from a dislocated sentence such as *As for the President I saw him in Chicago a few days ago* we cannot naturally form **In Chicago as for the President I saw him a few days ago*. Such examples suggest that it is difficult for a sentence to present more than one marked topic.

It is, however, fully natural to topicalize from an already passive sentence. Thus from *The President was welcomed with open arms in Chicago* we may naturally form *In Chicago the President was welcomed with open arms*. It appears then that the foregrounding inherent in passives does not compete with that expressed by topicalization or left-dislocation.

Moreover, the fact that we can topicalize or dislocate from a passive sentence is merely one example of a much broader difference in the syntactic nature of passive, compared with topicalization and dislocation. It is quite generally the case that the major syntactic operations in a language, such as nominalizing operations (*I was dismayed at John's being fired*), relative clause formation (*the garden in which John was attacked*), and yes/no question formation (*Was John attacked in the garden?*), operate freely on passives (with some exceptions, such as imperative formation), but these processes do not operate freely, often not at all, on topicalized or dislocated sentences. Thus we cannot say **I was dismayed at as for John his being fired*, or **the garden in which as for John Mary attacked him*, and so on. Generally then, basic passives tend to be well integrated into the rest of the grammar, whereas topicalizations and dislocations tend to be limited to main clauses, only sometimes being allowed in sentence complements of verbs of thinking and saying.

Furthermore, these basic differences between passives and topicalizations are directly reflected in the observable surface forms of passives. Consider how we can tell if a sentence in a language is passive or not. What is it about passives that makes them observably distinct in surface form from basic actives? For topicalizations and dislocations the informal answer is easy. They present NPs in 'unusual' positions in the sentence, that is, positions in which such NPs would

not occur in basic actives. In addition, in some languages, Lisu (Tibeto-Burman) and Japanese for example, these NPs may carry a specific marker of topichood, such as a postposition.

But passives are not in general distinct from actives with regard to the position and case marking of NPs. In particular the foregrounded NP in a passive, namely the derived subject, is usually placed and case-marked as are subjects of basic actives. Similarly, 'agent phrases', such as *by Mary* in *John was slapped by Mary*, most commonly take the position and case marking (including choice of pre- and postpositions) of some oblique NPs in active sentences, most usually an instrumental, locative, or genitive. Thus we cannot recognize a passive in terms of its NPs being marked or positioned in the sentence in ways different from those used in basic actives.

Note in particular that this holds for those languages which place the subject at the end in basic actives (see Keenan (1978) for a more extensive discussion). Thus in Kiribatese (Micronesia) the basic active order is Verb + Object (if present) + Subject + Oblique NP. And in the passive, (4b) below, the derived subject is placed where subjects of intransitive verbs normally occur in actives, and the agent phrase, constructed with a preposition, occurs where obliques normally go (the subscripts indicate agreement on the verb):

(4) a. E_i kamate-a_j te naeta$_j$ te moa$_i$
        it kill-it   the snake the chicken
        'The chicken killed the snake'

     b. E_j kamate-aki te naeta$_j$ (iroun te moa$_i$)
        it kill-PASS   the snake (by   the chicken)
        'The snake was killed (by the chicken)'

In fact the only way we know that (4b) above is passive is by the presence of a specifically passive suffix, *-aki*, on the verb. And this observation turns out to be general across languages.

That is, in general in a language, what is distinctive about the observable form of passives is localized within the predicate or verb phrase (understood broadly enough to cover auxiliary verbs). By contrast, topicalizations and dislocations are not generally marked in the predicate; the VPs in the topicalized and dislocated sentences cited above are identical to the VPs in their untopicalized and undislocated versions. Thus the formation of passives in a language takes place at the level of verb-phrase syntax, whereas topicalization and left-dislocation (as well as right-dislocation: *He's out of work again, my father*) take place at the level of sentence syntax. Stated in generative terms, to form a passive sentence it is sufficient to generate a passive verb phrase; the rules which combine these VPs with NPs to form sentences are rules needed for the formation of simple actives anyway and are not peculiar to passive. In contrast, the rules needed to form topicalizations or dislocations will derive sentences from sentences, and

will crucially refer to properties of the sentence as a whole, since they must specify the position to which the topicalized or dislocated element is moved *with respect to the sentence as a whole* - i.e., it is moved to the *front of the sentence* (or to the back in the case of right-dislocations).

Consequently, in examining passives in different languages, one should look for ways of forming verb phrases, not ways of modifying sentences to yield other sentences. And it is this point of view which we adopt in section 2 below in representing the language-general properties of passives.

We might conclude this section by emphasizing that the distinction between sentence-level phenomena and predicate-level ones is deeper and more extensive than simply a difference among foregrounding operations. Thus, if passive is thought of as a way of deriving sentences from sentences, as was the case in early forms of generative grammar (Chomsky (1957)), we would expect that, given a sufficiently large sample of languages, any of the ways in which one sentence could be derived from another would be used in the formation of passives in one or another language. But in fact this is very much not the case. Contrast passive with the formation of yes/no questions, clearly on all accounts a sentence- (or clause-)level derivational process. There are basically two major (not exclusive) means of forming such questions: beginning with a declarative, assign the declarative a distinctively interrogative intonation contour; or insert a particle, where the position of the particle is defined with respect to the declarative sentence as a whole, usually at the beginning of the sentence or at the end, more rarely between the subject and the predicate or after the first word or constituent of the sentence. Even such uncommon ways of forming questions as inverting the subject and the predicate or auxiliary verb are essentially sentence-level phenomena, as the smallest linguistic unit which contains the elements mentioned is the sentence. Thus what is distinctive about the observable form of yes/no questions is given by describing properties of the sentence as a whole (intonation contour, position of particle, etc.).

But *passives are never formed in such ways*. No language forms passive sentences by assigning a characteristic intonation contour to an active, or by inserting a sentence-level particle in an active, or by inverting the subject and the auxiliary of an active. Rather, passives are formed by deriving verb phrases in certain ways, ways to which we now turn.

2   Basic passives

2.1   General properties of basic passives

We shall refer to passives like (1b), *John was slapped*, as 'basic passives'. What makes them distinct from other passives is (i) no agent phrase (e.g. *by*

Mary) is present, (ii) the main verb in its non-passive form is transitive, and (iii) the main verb expresses an action, taking agent subjects and patient objects in its non-passive form. Our justification for calling such passives 'basic' is that they are the most wide-spread across the world's languages. More specifically, let us note the following generalizations concerning the distribution of passives:

G-1: Some languages have no passives.
G-2: If a language has any passives it has ones characterized as basic above; moreover, it may have only basic passives.

In support of G-1 we note that many languages in New Guinea, like Enga (C. N. Li and Lang (1979)), are cited as having no passives. Similarly, Chadic languages are typically passiveless (Hausa being a partial exception here; see Jaggar (1981)). Also passiveless are Tamang (Sino-Tibetan; Mazaudon (1976)), Isthmus Zapotec (Oto-Manguean; Pickett (1960)), and Yidiŋ (Australian; Dixon (1977a).

One might wonder whether these languages have a gap in their expressive power. Can they not express 'John was slapped' without committal as to who the agent was? And of course in general they can, but they will use fully active means to do so. If English had no passive, for example, we might give an approximate semantic equivalent by saying *someone slapped John*. It appears, however, that languages without passives have somewhat more grammaticized means for expressing functional equivalents of basic passives. Perhaps the most common means is to use an active sentence with an 'impersonal' third person plural subject. By impersonal here we mean simply that the third person element is not understood to refer to any specific group of individuals. Example (5b) below from Kru (John Singler (personal communication)) is illustrative:

(5)   a. Tò   pō   slā   ná
           Toe build house DEF
           'Toe built the house'

        b. Ī   pō   slā   ná
           3PL build house DEF
           'They built the house' = 'The house was built'

The functional equivalent to passive is often used in languages which have fully productive basic passives. Example (6) from Hebrew is illustrative.

(6)   Ganvu   li   et   ha-mexonit
        stole(3PL) to.me DO the-car
        'They stole my car' = 'My car was stolen'

A second alternative to passives is simply to eliminate the subject of the active; compare the active sentence in (7a) from Supyire (Gur) (from Carlson (1994))

with the passive sentence in (7b); the fact that *sikàŋi* 'goat' precedes the perfective marker in (7b) shows that it is the subject.

(7) a. nàŋa   à   sikàŋi   bò
          man.DEF   PERF   goat.DEF kill
          'The man killed the goat'

     b. sikāŋa   a   bò
          goat.DEF PERF   kill
          'The goat has been killed'

(Note that the difference between the two forms for 'goat' in (7) is purely phonological: the final /a/ on *sikāŋa* is due to assimilation to the /a/ of the perfective marker.) This alternative appears to be particularly common in ergative languages, such as Tongan, as in (8).

(8) a. Na'e tamate'i 'e   'Tevita 'a   Koliate
          killed   ERG   David   ABS   Goliath
          'David killed Goliath'

     b. Na'e tamate'i 'a   Koliate
          killed   ABS   Goliath
          'Goliath was killed'

It is not clear whether we want to consider such cases as (8b) as 'truncated' actives, with perhaps a third person plural or indefinite pronoun understood or as some kind of morphologically degenerate passive in which the verb is not distinctively marked.

A third and less common alternative to passive is to use a form of the verb which indicates an indefinite or unspecified subject. This is illustrated by the Oneida example in (9) (Iroquoian; Karin Michelson (personal communication)) in which the prefix *ukw-* is unambiguously a pronominal morpheme rather than a passive morpheme.

(9) úhkaʔ ok   waʔ-ukw-alahsʌtho-ʔ
     PRT   PRT   FACTUAL-UNSPEC.SUBJ:1.OBJ-kick-PUNCT
     'Someone kicked me'

This sentence serves the same function as a passive ('I was kicked'); however, it is not passive, but is active and transitive (like the English gloss 'Someone kicked me'). We discuss such constructions in section 4.2 below.

Consider now the distributional claim made in G-2. As formulated, it entails

G-2.1 to G-2.3 below:

G-2.1 If a language has passives with agent phrases then it has them without agent phrases.

G-2.2 If a language has passives of stative verbs (e.g. *lack*, *have*, etc.) then it has passives of verbs denoting events.

G-2.3 If a language has passives of intransitive verbs then it has passives of transitive verbs.

G-2.1 is not surprising, since agent phrases in passives are typically presented like oblique NPs in actives, and obliques are generally not obligatory. We should note here that Lawler (1977) cites Acehnese (Indonesia; Austronesian) as having a passive construction requiring an agent phrase. Durie (1987), however, argues with additional data that the construction is in fact an unmarked active. Conversely, many languages are cited as permitting only agentless passives; Latvian (see Lazdina (1966), from which (10) below is taken) is one example:

(10) Es tieku macits (*no mates)
      I   am   taught by mother
      'I am taught'

Similarly, contrast the active sentence from Taba (Indonesia; Austronesian; Bowden (1997)) in (11a) with the passive sentence in (11b), in which there is no expression of the agent.

(11) a. i   n=bes   niwi
          3sg 3sg=husk coconut
          'She husked the coconut'

      b. niwi   ta-bhes   do
         coconut NO.AGENT-husk REALIS
         'The coconut has been husked'

The prefix indicating passive in (11b) is in fact more generally an indicator of the absence of an agent; hence the gloss 'NO.AGENT'. With semantically transitive verbs, it serves as a passive marker, signalling that the sole argument of the verb corresponds to the patient in a corresponding active clause. But this prefix can also be used with intransitive verbs to indicate diminished agency on the part of the single argument, as in (12).

(12) ta-tagil   yak
      NO.AGENT-walk 1sg
      'I'm wandering around (with no specific destination in mind)'

And the passive construction in (13) from Kutenai (isolate; western Canada, USA) cannot include any reference to the agent.

(13) ɬa   c̓inamnaɬ-iɬ-ni   ʔinɬak   ʔaˑkitɬanamis
      back take-PASS-INDIC   chicken.hawk tent
      'Chicken Hawk was taken back to the tent'

In addition, it is also generally the case that agentless passives are pre-ferred even when the language syntactically permits agent phrases. Passive sentences with agent phrases are often accepted by native speakers of vari-ous languages (e.g. Turkish) with reluctance, and they are often described as reflecting the influence of English. And text counts for various languages (e.g. English by Svartvik (1966), Dutch by Kirsner (1976), Chamorro by Coore-man (1987), Modern Greek by Roland (1994)) show agented passives as much less frequent than agentless ones, even though agented ones are fully grammatical.

Regarding G-2.2, it should be noted that passives are often not formed freely on transitive verbs whose objects are not patients, not portrayed as being affected. Thus English verbs such as *be*, *become*, *lack* and *have* (in its pos-sessive sense, e.g. *John has a new car*) do not easily passivize (**A new car is had by John*). On the other hand, it is not the case, as has sometimes been sug-gested, that highly stative verbs are universally unpassivizable. In Kinyarwanda (Bantu, from Kimenyi (1980)), such highly stative verbs as *cost*, *weigh*, and possessive *have* do passivize:

(14)   Ibifuungo bibiri bi-fit-w-e   n-îshaâti
        buttons   two   they-have-PASS-ASP by-shirt
        'Two buttons are had by the shirt'

As regards G-2.3, we note (and discuss in detail in section 3.2) that many languages with basic passives allow the passive morphology to apply to intran-sitive verbs as well. For example, just as from *amare* 'to love' in Latin we form *amatur* 'he is loved', from *currere* 'to run' we form *curritur* 'it is run' in the sense 'there is running going on, running is being done'. And G-2.3 guarantees that if passives of the *curritur* type are present then so are passives of the *amatur* type. On the other hand, some languages with basic passives, like English, do not permit passives on intransitives. Passives on intransitives are the clearest examples of passives which lack the property of prototypical passives in which the subject corresponds to an object in a corresponding active clause. They do, however, entail the existence of an agent. The reason for defining passive so as to include such passives on intransitives is that they generally employ the same morphology as that used with basic passives and they normally eliminate an argument, the agent.

2.2   *The syntactic form of basic passives*

In section 1 we noted that what is distinctive about the form of passive sentences is their verb phrase (VP), and passive VPs are naturally expressed in the simplest case as syntactic and morphological modifications of transi-tive verbs (TVs). More specifically, a passive VP in a language will consist

of a strict morphological modification of a TV together with, in some languages, an auxiliary verb specific to the passive construction. This characterization of passives allows us to distinguish two broad types of passives: those which use auxiliaries, which we shall call 'periphrastic passives' and those which don't, which we shall call 'strict morphological passives'. The latter have already been illustrated by many examples cited above. An example of a passive with auxiliary, other than the English case, is the Latin example in (15).

(15)   Dareus (ab Alexandro) victus   est
           Darius (by Alexander) conquered is
           'Darius was conquered (by Alexander)'

This example shows in addition that Latin in fact possesses passives of both types, since it also possesses the morphological passive mentioned above, illustrated by *amatur*. As we shall see below, it is quite common for a language to have more than one syntactically and semantically distinct type of passive construction.

2.2.1   Strict morphological passives   The strict morphological (SM) passives illustrated so far are all formed by suffixing, but this of course is not a general property of SM-passives. Example (16b) from Sre (Mon-Khmer; Manley (1972)) illustrates a passive formed by prefixing.

(16)   a. Cal   paʔ mpon
          wind open door
          'The wind opened the door'

      b. Mpon gə-paʔ   mə cal
          door PASS-open by wind
          'The door was opened by the wind'

Examples of other morphological ways in which passives are formed include infixing (Tagalog), internal vowel change (Hebrew, Arabic), and reduplication (Hanis Coos, western USA).

  A given language may present several formally distinct SM-passives. Examples (17b–d) below, from Malagasy, are illustrative.

(17)   a. Man   + tsangana (= manangana) ny lai aho
          ACTIVE + put up   the tent I
          'I am putting up the tent'

      b. *A*-tsanga-ko   ny lai
          PASS-put.up-by.me the tent
          'The tent is put up by me'

    c. *Voa*-tsangana ny   lai
       PASS-put. up   the tent
       'The tent is put up'

    d. *Tafa* -tsangana ny   lai
       PASS-put.up   the tent
       'The tent is put up'

Formally, the three passives above differ with respect to the choice of passive prefix on the verb. In addition, the last two do not easily accept agent phrases, though their presence is perhaps not strictly ungrammatical. Semantically, the three passives are not fully equivalent, however. Example (17b) is a neutral passive, forming a paraphrase with the active in (17a). The *voa-* passive in (17c), however, is unequivocally perfective in meaning: the action of putting up the tent is viewed as successfully completed. The meaning of (17d) is somewhat harder to describe, but roughly it suggests that the action of putting up the tent was almost spontaneous; the conscious activity of the agent is downplayed. We might almost be tempted to translate (17d) as 'the tent put itself up', though of course that could not happen literally. See Randriamasimanana (1986) and Rajaona (1972) for more thorough discussion.

    The same types of formal morphological means used in deriving basic passives are often used to derive VPs which are not passives. This is particularly true for verbal morphology commonly associated with reflexives and/or middles. In Spanish, for example, the reflexive construction can be used as a passive, as in (18).

(18)   Se   encontraron dos nuev-o-s   cuadros   de Frida Kahlo
       REFL find.PAST.3PL two new-MASC-PL paintings.PL by Frida Kahlo
       'Two new paintings by Frida Kahlo were found'

A similar construction exists in Russian, and here it is also possible to express the agent phrase:

(19)   Doma   strojat-sja   rabočimi
       houses build-REFL workers.INSTR
       'Houses are built by workers'

    There are two further properties of SM-passives which should be noted. First, in some cases, the morphological function is 'degenerate' in that the derived expression does not differ at all from what it is derived from. Such cases of degenerate morphological functions are not uncommon. For example, the function which forms past participles in English (*kicked* from *kick*, *eaten* from *eat*, etc.) is degenerate in certain cases; the past participle of *hit* is simply *hit*, not *hitted* or *hitten*. One might expect, then, to find passive VPs which are identical to the transitive verbs they are derived from. And some few cases

Table 6.1 *Conjugation of Latin* amare

	Present indicative active		Present indicative passive	
	singular	*plural*	*singular*	*plural*
1	amo	amamus	amor	amamur
2	amas	amatis	amaris	amamini
3	amat	amant	amatur	amantur

seem to exist, though they are not common and usually of restricted distribution in the languages for which we have data. Thus the verb in (20b) from Swahili (Givón (1972)) does not differ from its active transitive counterpart in (20a) (except that it shows subject agreement with an NP in a different noun class):

(20) a. Maji   ya-meenea nchi
          water it-cover   land
          'The water covers the land'

      b. Nchi   i-meenea   maji
          land   it-cover   water
          'The land is covered by water'

Similar examples are cited for other Bantu languages, e.g. Kinyarwanda (Kimenyi (1980)). Kimenyi in particular notes a very large number of constraints both on the formation of such passives and on their distribution in various syntactic contexts.

   Finally, given that SM-passives are derived VPs, it is always possible that other syntactic or morphological processes which operate on VPs may be sensitive as to whether the VP in question is passive or not. We shall illustrate this possibility here with the case of verb (more exactly verb phrase) agreement with subjects. The main point here is that the existence and form of subject agreement affixes on passive verbs may differ from those on active verbs. In (i–iii) below we give the principal types of such variation known to us:

   (i) *The passive verb may fail to agree with its subject, even though actives do show agreement*. This is the case in Welsh for the SM-passive (there is also a periphrastic passive). Active verbs agree with pronominal subjects, but passive verbs remain invariant: *gwelir di* 'You are seen', *gwelir fi'* 'I am seen', *gwelir ef* 'he is seen', etc.

   (ii) More commonly than (i) above, *passive verbs may simply have different agreement affixes from active verbs*. This is the case for example with the SM-passive in Latin. Compare the present indicative actives of *amare* 'to love' with their present indicative passives in Table 6.1. Clearly the variation in

person and number in the passive forms is not identical to that of the actives. It is quite common across languages that agreement forms may vary with the other properties which are marked on the verb. Thus person–number endings on verbs in Romance languages may vary with tense, mood and aspect. So the variation noted above is to be expected as long as passive is a verbal category, not a sentential one. If passive were merely thought of as an operation which topicalized an NP and perhaps backgrounded another, we might expect markings of passive to show up on NPs, but not on the VPs. And in particular we would have no reason to expect that verbs in such 'topicalized' sentences would show different agreement paradigms from their non-topicalized (active) counterparts.

(iii) *The passive verb may agree with its subject as though it were a direct object of an active verb*. This is the case in Maasai (Nilo-Saharan) and Kimbundu (Bantu; Angola). Example (21) below is from Kimbundu:

(21)   a. A-mu-mono
            they-him-saw
            'They saw him'

        b. Nzua a-mu-mono   kwa meme
            John they-him-saw by   me
            'John was seen by me'

The sentence in (21b) qualifies as a passive to the extent that the patient is in subject position before the verb and the agent is expressed in a prepositional phrase following the verb. But the verb exhibits semantically empty third person plural agreement and object agreement with the patient. (It is tempting to speculate in this latter case that the passive in (21b) derives historically from an object topicalization from an impersonal third plural active of the sort illustrated in (5b) and (6).)

2.2.2   Periphrastic passives   A basic periphrastic passive consists of an auxiliary verb plus a strict morphological function of a transitive verb. These passives fall into natural classes according to the choice of auxiliary verb; the passive auxiliary can be (i) a verb of being or becoming; (ii) a verb of reception; (iii) a verb of motion; or (iv) a verb of experiencing. We elaborate on each of these four types in the following paragraphs.

(i) *The auxiliary verb is a verb of being or becoming*. Example (22) below from German and (23) from Persian illustrate the use of 'become' as a passive auxiliary:

(22)   Hans wurde   von seinem Vater   bestraft
        Hans became 'by' his   father   punished
        'Hans was punished by his father'

(23)   a. Ali Ahmed-ra   košt
          Ali Ahmed-OBJ killed
          'Ali killed Ahmed.'

       b. Ahmed košté šod
          Ahmed killed become
          'Ahmed was killed'

The use of 'be' as an auxiliary is illustrated by the standard English passive, *John was slapped*, as well as by (15) from Latin and (10) from Latvian. Note that, in several of these cases, the verb form with which the auxiliary combines is a 'past participle', a form that in some ways behaves like an adjective: for example, it agrees in Latin and Russian with the subject of the passive VP in number and gender but not person, which is the agreement paradigm for adjectives rather than verbs.

It should be noted as well that periphrastic passives of the 'be' sort commonly exhibit a certain ambiguity (or vagueness) as to whether they are interpreted 'dynamically' or purely 'statively'. Thus *The vase was broken* is ambiguous in English as to whether it merely specifies a state of the vase (which might in fact not have been caused by an external agent) or an activity performed upon the vase. German, in contrast, avoids this kind of ambiguity and always permits two different structures for these cases:

(24)   a. Das Haus   wird   verkauft
          the   house becomes sold
          'The house is being sold'

       b. Das Haus   ist verkauft
          the   house is   sold
          'The house is sold'

If (24a) obtains you will have a chance to buy the house, whereas if (24b) obtains you are too late. Note that 'get' passives in English (e.g. *The vase got broken*, *John got fired*, etc.) have only the dynamic interpretation.

(ii) *The passive auxiliary is a verb of reception* (e.g. *get*, *receive* or even *eat*). In such cases it is common that the modification of the transitive verb takes the form of a nominalization, that is, something which occurs independently in the language as a nominal form of some sort. Example (25) below from Welsh and (26) from Tzeltal (Mayan) illustrate this type:

(25)   Cafodd Wyn ei   rybuddio gan Ifor
        get   Wyn his warning   by   Ifor
        'Wyn was warned by Ifor'

(26)   La   y-ich'   'utel   (yu'un   s-tat)   te   Ziak-e
        PAST he-receive bawling.out (because his-father) ART Ziak-ART
        'Ziak got a bawling out (from his father)'

In fact English constructions like *John got a licking / tongue lashing / beating from Bill* appear to illustrate this sort of passive, though they are highly limited as to which transitive verbs accept them; we cannot say **Bill got a killing/praising/etc., from Harry*. English 'get' passives (e.g. *John got killed*) resemble passives of this sort, but differ in that the thing got does not take the form of a nominalization. Furthermore, the direct source of 'get' passives seems to be the inchoative sense rather than the original sense of 'receive'. Thus, *John got killed* is parallel to *John got sleepy*.

   (iii) *The passive auxiliary is a verb of motion* (e.g. *go*, *come*). This type seems less well attested than either (i) or (ii) above, but examples (27) below from Hindi and (28) from Persian are suggestive:

(27)   Murgi   mari   gayee
        chicken killed went
        'The chicken was killed'

(28)   a. Ali loget-ra   be kar   bord
           Ali word-DO   to work   take
           'Ali used the word'

        b. Loget be kar   reft
           word to work   went
           'The word was used'

   (iv) *The passive auxiliary is a verb of experiencing* (e.g. *suffer*, *touch*, even '*experience pleasantly*'). Example (29) below from Thai and (30) from Vietnamese (diacritics omitted) are illustrative:

(29)   Mary   thúuk   (John)   kóot
        Mary   touch   (John)   embrace
        'Mary was embraced (by John)'

(30)   Quang   bi   (Bao)   ghet
        Quang   suffer (Bao)   detest
        'Quang is detested (by Bao)'

   Passives of this sort are widely attested in languages spoken in southeast Asia, including Mandarin, although their analysis as passives is in fact not obvious. The languages which exhibit them are independently verb-serializing languages: apparently simplex sentences are commonly constructed with multiple verbs and few if any prepositions. For example, a sentence such as 'John took the train to Boston' might be literally rendered as 'John go ride train arrive

Boston'. In addition such languages exhibit virtually no bound morphology. And since passive auxiliaries can quite generally (but not always) occur as main verbs in simple sentences, it is plausible to analyse passive sentences in these languages as special cases of serial-verb constructions. We refer the reader for further discussion of these constructions in Vietnamese to Nguyen (1976), where reasonable evidence is given that the verb of experiencing is functioning as an auxiliary verb.

Accepting these structures as passives, it should be noted that there will commonly be several acceptable choices for passive auxiliaries. Nguyen (1976) cites five such verbs for Vietnamese, among them *duoc*, used when the subject is portrayed as pleasantly affected by the action:

(31)   Quang duoc   Bao thuong
        Quang 'enjoy' Bao love
        'Quang is loved by Bao'

The use of the auxiliary *bi* 'suffer' is possible in (31) but ironic.

2.3   The semantics of basic passives

Our discussion of passives in section 1 in terms of their corresponding actives suggests that we might think of passives simply as paraphrases of their corresponding actives. Ultimately, however, such a view is mistaken. For one thing, as already discussed, basic passives lack an agent phrase and thus lack a corresponding active, strictly speaking. For another, as we will see shortly, when the NPs denote things other than individuals, the passives are often not paraphrases of their corresponding actives. On the other hand, if we view passive as an operation on active transitive verb phrases (TVPs), deriving passive VPs, we can, as a first approximation, give the interpretation of passives as follows: the passive of a TVP is true of an individual x if and only if, for some individual y, the TVP is true of the pair (y, x). So *was slapped* holds of John if and only if, for some individual y, *slap* holds of $(y,$ John$)$; that is, if someone slapped John. Notice that this semantic interpretation makes no immediate claim concerning whether passive *sentences* are paraphrases of their actives. Given that NPs like *John* and *Mary* denote individuals, it will claim that *Mary slapped John* entails *John was slapped*. But if the NPs in the sentence are not the sort which denote individuals, then no such entailment paradigm regularly holds. Thus *No student slapped John* will not entail *John was slapped*. Nor will it be the case that passive sentences in general entail the corresponding existential generalization of the active. Thus *Every cake was stolen* does not entail that some individual x stole every cake. Different cakes might have been stolen by different individuals.

The main point we want to notice here, however, is that if passives are treated as ways of deriving VPs from TVPs we can give a basically correct semantic interpretation which is in accordance with the following general principle: *the semantic interpretation of derived structures depends on (is a function of) the meanings of what they are derived from.* On the other hand, if passive were thought of as a way of deriving sentences from sentences, no regular semantic relationship between the derived structure and what it is derived from could be given. Sometimes an agentless passive is entailed by an active and sometimes it isn't, as we have seen above. Moreover, the same disparity exists between agented passives and their actives. Thus, while *John was kissed by Mary* is presumably logically equivalent to (has the same truth conditions as) *Mary kissed John*, a sentence like *Each child was kissed by no politician* is clearly not logically equivalent to *No politician kissed each child*. Thus, treating passives as a VP derivational process correctly allows us to predict that quantified NPs in passives and actives may exhibit different scopes.

We turn now to some further generalizations surrounding the semantics of passives.

G-3: Languages with basic passives commonly have more than one formally distinct passive construction.

Moreover, distinct passives in a language are likely to differ semantically with respect to aspect and/or degree of subject affectedness, some examples of which we now turn to.

2.3.1   Aspectual differences

G-4: If a language has any passives it has ones which can be used to cover the perfective range of meaning.

G-5: If a language has two or more basic passives they are likely to differ semantically with respect to the aspect ranges they cover.

From G-4 we may infer that languages like Russian with a specifically imperfective passive will also present a passive construction which covers the perfective range. Thus no language will have only passives which must be interpreted imperfectively.

G-5 has already been amply illustrated. Recall the three basic passives in Malagasy (example (17)), one of which was semantically rather neutral, the second (the *voa-* passive) clearly perfective, and the third (the *tafa-* passive) indicating something like spontaneous action, with little intentional involvement of an agent. Recall as well that in languages like Russian, Latin, and Kinyarwanda with a strict morphological passive and also a 'be' type periphrastic

passive, the periphrastic form is commonly interpreted as stative or perfective with respect to the SM-passive which is either non-committal as to aspect or else specifically imperfective. And recall finally the distinction between dynamic passives, which focus attention on the action, as opposed to stative passives which focus attention on the state of the object, perhaps regardless of whether an external agent is responsible. Thus the 'get' passives in English are dynamic and the 'be' passives at least ambiguously stative. A similar distinction among SM-passives is illustrated in (32) below from K'ekchi (Mayan):

(32) a. Laʔin sh-in-sakʔ-eʔ
         I   PAST-1(ABS)-hit-PASS
         'I was hit [emphasizes action of hitting]'

      b. Laʔin sakʔ-bʔil-in
         I   hit-PASS-1(ABS)
         'I am the one who is hit [emphasizes the resultant state]'
                                 Ava Berenstein (personal communication)

2.3.2 Degree of subject affectedness We note the following generalizations:

G-6: The subject of a passive VP is always understood to be as affected by the action as when it is presented as the object of an active transitive verb.

G-7: Distinct passives in a language may vary according to degree of affectedness of the subject and whether it is positively or negatively affected, though this variation seems less widely distributed than that of aspect.

Recall in these regards the *bi* versus *duoc* passives in Vietnamese, (30) and (31), in which the subject of a *bi* passive is understood to be negatively affected, whereas the subject of a *duoc* passive is understood to be positively affected. In addition, Vietnamese may use the passive auxiliary *do* which is semantically neutral as regards subject affectedness, as in (33):

(33) Thuoc   X do   Y che   nam   1973
     medicine X PASS Y invent year 1973
     'Medicine X was invented by Y in 1973'

Negative effect passives (often called 'adversative passives') seem to be the norm in much of eastern Asia. Thus the common *bei* passive in Mandarin is often interpreted as negatively affecting the subject; similarly for the standard *-are-* passive in Japanese. And as regards Korean, C. M. Lee (1973) cites, in addition to the *-ki-* passive, a negative effect passive constructed with *tangha* 'to be subjected to', illustrated in (34), and a positive effect passive constructed with *pat* 'to receive' as in (35):

(34) Pʰoro-ka   henpjeng-eke kutʰa-tangha-ess-ta
      POW-SUBJ   MP-AG   beat-subject.to-PAST-DECL
      'The prisoner of war was beaten (subjected to a beating) by an MP'

(35) Kɨ sensayng-ɨ n haksayng-tɨ l-eke conkjeng-pat-ɨ n-ta
      the teacher-TOP   student-PL-AG   respect-receive-PRES-DECL
      'The teacher is respected by students'

3   Non-basic passives

In section 2 we considered the syntactic and semantic properties of basic pas-
sives, which were defined as ones which lacked agent phrases and were formed
from transitive verbs denoting events. In this section, we turn to a variety of
non-basic passives. We will first consider passives that are non-basic by virtue of
including an agent phrase, and then turn to a variety of less common non-basic
passives formed from verb phrases that are not simple transitive verb phrases.

3.1   *Passives with agent phrases*

We consider passive clauses that include an agent phrase to be non-basic. There
are three reasons for considering that agent phrases are not in general an integral
part of the passive construction itself: (i) many languages present passives which
do not permit agent phrases; (ii) agent phrases occur in non-passive structures;
and (iii) when present, agent phrases most commonly take the form of an
independently existing oblique NP in actives. We consider (ii) and (iii) in this
section.

   Let us consider first somewhat more closely what is meant by 'agent phrase'.
To say that *by Mary* is the agent phrase of *John was kissed by Mary* is to say
that *Mary* functions as the semantic subject but not the syntactic subject of the
transitive verb *kiss*, from which the passive VP is derived. In general, an agent
phrase is an NP (with or without adpositions) which functions as the semantic
but not syntactic subject of a verb in an expression derived from that verb (or
verb phrase). Note that the term 'agent phrase' is potentially misleading in that
its semantic role (agent, experiencer, etc.) is whatever is required by the verb
of which it is the understood subject, and need not be specifically agent, as in
the example *Money is needed by the church*. Some linguists use the alternative
term 'actor' to emphasize this point.

3.1.1   *Agent phrases in non-passive constructions*   It is important to note
first that agent phrases can often occur in structures which are not passives.
Consider for example the *-ing* nominals in English in (36):

(36)   a. The university forbids talking by students during exams
         b. Cheating by students is punishable by expulsion

Clearly *students* is the semantic (but not syntactic) subject of the intransitive
verbs *talk* and *cheat* in these examples. *By students* then is an agent phrase. And
since intransitive verbs in English do not passivize, we infer that agent phrases
occur independently of the passive construction in English.

   A second non-passive construction in which agent phrases appear is the
causative (especially the indirect as opposed to direct causative). The examples
below, (37a) from Japanese (Howard and Niyekawa-Howard (1976)) and (37b)
from German, are illustrative:

(37)   a. Zyon ga   Biru ni aruk-ase-ta
            John SUBJ Bill by walk-CAUS-PAST
            'John had Bill walk'

         b. Seine erste Frau liess sich von ihm scheiden
            his   first wife let self by him divorced
            'His first wife let herself be divorced by him'

In both cases in (37), the agent phrase is marked by the same adposition that
is used for agent phrases in passives, though in causatives generally the choice
of case or adposition for the understood subject of the causativized verb may
vary as a function of the transitivity of the underlying verb. Nonetheless, the
marking of the agent phrase as in passives shows up with more than chance
frequency.

3.1.2   The form of agent phrases   Most commonly, agent phrases, whether
in passives or other constructions, are presented as (i) instrumentals, (ii) loca-
tives, or (iii) genitives, in active constructions.

   (i) *Instrumentals*. Example (19) above from Russian illustrates the use of
the instrumental case marking on agent phrases, the same case as is used of
course for instruments, as in *John cut the bread with a knife*. Similarly, in many
Bantu languages the agent phrase is marked with the preposition *na* which
independently marks instruments in actives, as illustrated in (38) below from
Kinyarwanda; (14) above illustrates a passive with an agent phrase marked by
the same preposition:

(38)   Umugabo araandika ibaruwa n-ikaramu
         man   write   letter with-pen
         'The man is writing a letter with a pen'

(ii) *Locatives*. The agent phrase marker in English is independently used with locative force: *He sat by the window and was seen by Mary*. More common, however, is for the agent phrase marker to be specifically the ablative marker (the locative marker meaning 'from'), as in Kayardild (Tangkic; Australia; Evans (1995)):

(39)   ngada   ra-yii-ju   mun-da   balarr-ina maku-na
         1SG.NOM bite-MID-POTENT buttock-NOM white-ABL woman-ABL
         'I will be injected in the buttocks by the white woman'

(iii) *Genitives*. The presentation of agent phrases as possessors in Malagasy is illustrated in (40).

(40)   a. ny   entan-dRakoto
             the packages-Rakoto
             'the packages of Rakoto'

         b. Nosasan-dRakoto   ny   lamba
             wash.PASS-by.Rakoto   the clothes
             'The clothes were washed by Rakoto'

Similarly, both *von* in German and *de* in French have possessive uses *(ein Freund von mir* 'a friend of mine' and *un ami de Pierre* 'a friend of Pierre') and they are the agent markers in passive clauses (*Er wurde von Marie geküsst* 'He was by Mary kissed'; *Il est aimé de ses parents* 'He is loved by his relatives'), though they also have ablative uses (*Er fährt von Stuttgart nach Köln* 'He goes from Stuttgart to Cologne'; *Il vient de Paris* 'He comes from Paris').

There is, however, a minority of cases where agent phrases are not presented as instrumentals, locatives, or genitives:

(iv) *The agent phrase has no adposition*. Vietnamese (30) and (31) illustrate this case. Note equally example (41) from Haya (Bantu):

(41)   Ebitooke bí-ka-cumb-w'   ómukâzi
         bananas   they-PAST-cook-PASS woman
         'The bananas were cooked by the woman'
                                                        Duranti and Byarushengo (1977)

Note that such examples do not obviously violate the claim that agent phrases are presented like active obliques. Cross-linguistically, many nonsubjects and nonobjects are commonly presented with no adposition. This is especially common for temporals, as in *John saw Mary last week*.

(v) *The agent phrase is incorporated into the passive verb*. English illustrates this case for a very limited range of verbs, roughly a subset of those expressing power and authority:

(42)   a. This project is state-controlled/NSF-funded/government-
             regulated
         b. *This project is state-enjoyed/NSF-avoided/government-rejected

Such incorporation seems more productive in Quechua; in (43b), the agent apparently forms a close unit with the verb:

(43)   a. Kuru-ø   manzana-ta miku-rqa-n
             bug-SUBJ   apple-DO   eat-PAST-3
             'The bug ate the apple'

         b. Kuru miku-sqa-mi   manzana-ø ka-rqa-n
             bug   eat-PTCPL-COMMENT apple-SUBJ be-PAST-3
             'The apple was bug eaten'

   Finally, we should note that there may be a few cases in which agent phrases are introduced by an adposition which does not occur independently in oblique NPs in active structures. For example, the agentive preposition *al yedei* in Hebrew is largely limited to agent phrases in passives (though it is closely related to an active oblique preposition *al yad* 'near'). Similarly, the agent preposition *oleh* in Indonesian appears limited to agent phrases in passives.

3.2   *Passives on non-transitive verbs*

The notion of passive we have characterized so far seems very dependent on the notions of intransitive and transitive verbs. In fact, however, our notion naturally generalizes in ways which are linguistically enlightening, as they suggest the existence of passives different from those already considered, and languages do in fact present such passives. To see the generalization, let us replace the linguistic notion of 'verb phrase' with its logical counterpart, that of a 'one-place predicate phrase', namely something which combines with *one* NP to form a sentence. Similarly, the notion of 'transitive verb phrase' may be replaced by that of a 'two-place predicate phrase', something which combines with *two* NPs to form a sentence. In general an *n*-place predicate phrase will combine with *n* NPs to form a sentence. So the standard ditransitive verbs, *give*, *hand*, etc., might be considered 'three-place predicate phrases'.

   In these terms, the standard passive derives a one-place predicate (phrase) from a two-place predicate (phrase) (we henceforth drop the term 'phrase'). Generalizing over the number of NPs a predicate needs to form a sentence then, we may characterize passive as a way of deriving *n*-place predicates from $n + 1$-place predicates. The case we have treated so far is that for which *n* is 1. Let us consider now the case for $n = 0$. That is, do we find passives in languages which derive zero-place predicates (sentences) from one-place predicates (VPs)? And

the answer obviously is 'yes'. We have in fact already cited the Latin example whereby, from the one-place predicate *currit* 'is running', we derive the zero-place predicate *curritur* '(it) is run, running is being done'.

As passives on intransitives in the simplest cases will be lacking any NPs, they will of necessity be subjectless, and as such have usually been called 'impersonal passives' in the literature. While their general properties are less well known than the personal passives, we do have several studies that make remarks on a variety of languages and on which the remarks below are based. These studies are Comrie (1977), Perlmutter (1978), and to a lesser extent Keenan (1976b). For studies of specific languages we cite: Langacker (1976) for Uto-Aztecan languages generally, Kirsner (1976) for Dutch, Timberlake (1976) for North Russian dialects, and Noonan (1994) for Irish.

Based on these studies, which cover many fewer languages than those considered in our discussion of basic passives, we somewhat tentatively suggest the following general properties of impersonal passives (passives of intransitives).

First, such passives exist and seem to have a reasonable distribution across language areas and genetic families. Thus, languages such as the following have basic passives and use the same syntactic and morphological means to derive impersonal passives (sentences) from intransitive verb phrases: Dutch, German, Latin, Classical Greek, North Russian dialects, Shona (Bantu), Turkish, and Taramahua (Uto-Aztecan). For example, in German, basic passives are formed from 'become' plus the past participle of the transitive verb, and impersonal passives are formed from 'become' plus the past participle of the underlying intransitive verb, as in (44):

(44)   Gestern   wurde   getanzt
        yesterday became   danced
        'Yesterday there was dancing'

Similarly in the impersonal passive in (45a) below from Turkish we see the same *-il-* marking passive as in the basic passive illustrated in (45b) (the different form *-ıl* in (45b) is simply the result of vowel harmony).

(45)   a. Ankara-ya gid-il-di
            Ankara-to go-PASS-PAST
            'It was gone to Ankara' / 'There was a trip to Ankara'

        b. Pencere Hasan tarafından aç-ıl-dı
            window   Hasan by   open-PASS-PAST
            'The window was opened by Hasan'

And (46a, b) below from Taramahua clearly illustrate the same bound morphology on the verbs in each case:

(46)   a. Tashi goci-ru
          not   sleep-PASS
          'One doesn't sleep'

       b. Gao   ne   ʔa-ru
          horse I(SUBJ) give-PASS
          'I was given a horse'

Second, impersonal passives using the same verbal morphology as basic passives typically take their agent phrases marked in the same way as in basic passives, if they accept agent phrases at all. Example (47) from Dutch is illustrative:

(47)   Er   wordt   (door de jongens)   gefloten
       there became (by the young.men) whistled
       'There was some whistling by the young men'

Sometimes, as in Turkish, the passives on intransitives do not accept agent phrases whereas those from transitives do (though only with some awkwardness in Turkish). For most of the languages where we have data, however, as in Dutch, Latin, North Russian, and Shona, the impersonals accept agent phrases if the basic passives do.

Third, impersonal passives are not limited to lexical intransitive verbs. In particular, a transitive verb together with its NP object will constitute a VP and may be passivized; in this case, the NP will remain an object after passivization. Observe (48) from North Russian.

(48)   U mena bylo   telenka   zarezano
       at me   was(3SG.NEUT) calf(FEM.ACC) slaughtered(SG.NEUT)
       'By me there was slaughtered a calf'

Similarly, Latin employs impersonal passives with transitive verbs whose object occurs in the dative case, as illustrated in (49).

(49)   a. Boni cives   legibus   parent
          good citizens law(DAT.PL) obey
          'Good citizens obey laws'

       b. Legibus   (a bonis civibus) paretur
          law(DAT.PL) (by good citizens) is.obeyed(3SG)
          '(By good citizens) there is obeying laws'

Fourth, it appears in a few languages that 'reflexive' forms come to function with an impersonal passive meaning, as illustrated in (50) from Polish; note the impossibility of including an agent phrase.

(50)   a. Idzie   sie   szybko (*przez uczniow)
          is.walked REFL quickly by   schoolboys
          'One walks quickly'; 'There is quick walking'

       b. Dokonuje   sie   prace (*przez uczonych)
          is.completed REFL works by   scientists
          'The works are being completed'

And fifth, a language may have impersonal constructions which are syntactically and morphologically independent of the existence of basic passives. This is illustrated in (51) for Irish. In (51a) is an example of a simple active clause in Irish; (51b) is the basic passive of (51a); (51c) is an impersonal passive corresponding to (51a). Note that the impersonal passive construction in (51c) is completely different from the basic passive in (51b): in (51b), the basic passive is formed with an auxiliary plus participle, while in (51c), the impersonal passive involves a distinct form of the verb.

(51)   a. Bhuail si   e   (Active)
          hit   she him
          'She hit him'

       b. Bhi se buailte   aici   (Basic passive)
          AUX he hit(PTCPL) at.her
          'He was hit by her'

       c. Buaileadh (lei)   e   (Impersonal passive)
          hit(IMPRS) (with her) him
          'There was hitting of him (by her)'

In fact, mind-bogglingly, Noonan and Bavin-Woock (1978) shows that a basic passive in Irish may be further subject to the impersonal construction in that language: (52) is an impersonal passive formed on the basic passive in (51b).

(52)   Bhiothas   buailte   (aici)
       aux(IMPRS) hit(PART) (at.her)
       'There was being hit (by her).'

3.3   *Passives on ditransitive verb phrases*

Among the common three-place or ditransitive verb phrases in a language will be translations of verbs like *give*, *show*, *teach*, etc. Languages vary as to which of the two objects in such clauses can appear as the derived subject in a passive. In French, for example, only the patient and not the recipient can serve as the subject of a passive. Thus, given an active ditransitive clause like (53a), we find a passive with the patient as subject, as in (53b), but not one with the recipient as subject, as illustrated by the ungrammaticality of (53c).

(53)   a. Jean a   donné le livre à Pierre
            Jean has given the book to Pierre
            'Jean gave the book to Pierre'

        b. Le livre a   été   donné à Pierre
            the book has been given to Pierre
            'The book was given to Pierre'

        c. *Pierre a   été   donné le livre
            Pierre has been given the book
            'Pierre was given the book'

In Yindjibarndi, in contrast, only the recipient and not the patient of a ditransitive verb can become the subject of a passive clause. Thus, given an active clause like (54a), we can form a passive in which the recipient is subject, as in (54b), but (54c), in which the patient is subject, is ungrammatical.

(54)   a. Ngaara yungku-nha ngayu   murla-yi
            man   give-PAST   1SG.OBJ meat-OBJ
            'A man gave me the meat'

        b. Ngayi yungku-nguli-nha murla-yi   ngaarta-lu
            I   give-PASS-PAST   meat-OBJ   man-INSTR
            'I was given the meat by a man'

        c. *Murla yungku-nguli-nha ngayu   ngaarta-lu
            meat   give-PASS-PAST   1SG.OBJ   man-INSTR
            'The meat was given to me by a man'

In Kinyarwanda, on the other hand, both the recipient *and* the patient can serve as subject in a passive clause:

(55)   a. Umugabo y-a-haa-ye   umugóre igitabo
            man   he-PAST-give-ASP woman   book
            'The man gave the woman the book'

        b. Umugóre y-a-haa-w-e   igitabo n-ûmugabo
            woman   she-PAST-give-PASS-ASP book   by-man
            'The woman was given the book by the man'

        c. Igitabo cy-a-haa-w-e   umugóre n'ûmugabo
            book   it-PAST-give-PASS-ASP woman   by-man
            'The book was given to the woman by the man'

English, at first glance, is similar, but with complications. We find both recipient passives and patient passives:

(56)   a. John was given the book by Mary
        b. The book was given to John by Mary

Unlike the Kinyarwanda case, however, English also has two different active constructions for expressing such meanings:

(57)   a. Mary gave John the book
         b. Mary gave the book to John

The fact that for most speakers the preposition *to* is required in (56b) (**The book was given John by Mary*) suggests that (56b) is the passive of (57b) and that (56a) is the passive of (57a). There are a variety of proposals for describing the relationship between the two sentences in (57) (cf. Dryer (1986), and chapter 4, section 2.3) which we will not discuss here, but it appears that we can in general predict that if a language does exhibit two different active constructions analogous to those in (57), then the language will also exhibit two passives, analogous to those in (56).

In addition to the classical three-place predicates such as *give*, languages will commonly have three-place predicators like *put*, *place*, which require a patient and some sort of locative in addition to an agent to form a sentence. And again, the predicate will always have passive forms taking the patient as the subject if it has any passives at all. It may or may not have passives taking the locative as subject. English for example by and large does not have locative passives: **The chest was put the jewels (in)*. We may note that locative passives in some cases are easier if the verb is otherwise intransitive, e.g. *This bed was slept in*.

On the other hand, other languages, such as Kinyarwanda and Chichewa (and commonly Bantu), as well as Malagasy and the Philippine languages generally, quite productively form passives whose subjects are semantically locatives. There appear to be two general syntactic means for forming such passives. Moreover, these means are not specific to locatives but apply in general to passives which have non-patient subjects.

3.4   Other passives with non-patient subjects

One general means of forming such passives is first to modify the *n*-place predicate (*n* greater than 2) to a form in which the non-patient is treated like the (direct) object of simple transitive verbs, and then form a passive as is done generally in the language on transitive verb phrases. Example (58) from Kinyarwanda illustrates this strategy (from Kimenyi (1980)):

(58)   a. Úmwáalímu y-oohere-je igitabo kw-iishuûri
            teacher   he-send-ASP book   to-school
            'The teacher sent the book to school'

b. Úmwáalímu y-oohere-jé-ho ishuûri igitabo
   teacher   he-send-ASP-to school   book
   'The teacher sent the school the book'

c. Ishuûri ry-oohere-je-w-é-ho   igitabo n-úúmwáalímu
   school it-sent-ASP-PASS-ASP-to book   by-teacher
   'The school was sent the book by the teacher'

Note that in (58b) the basic verb 'send' is modified in form by the presence
of the goal locative suffix *-ho* (morphologically related to the goal locative
preposition *kwa)* and 'school' occurs immediately postverbally without a goal
locative marker. Kimenyi shows that this postverbal NP has the properties shared
by the sole objects of simple transitive verbs.

   In (58c) this complex verb is passivized in the normal way in Kinyarwanda,
by the non-final suffix *-w-*, and 'school' is clearly a derived subject, occurring
in subject position, triggering subject agreement on the verb, and in general
(as Kimenyi supports in detail) having the characteristic syntactic properties of
subjects of basic verbs. Bantu languages such as Kinyarwanda and Chichewa
(Trithart (1977)) are particularly rich in ways of presenting oblique NPs of
actives as derived objects. Essentially any oblique NP in Kinyarwanda can be
the surface direct object of some derived form of a verb. See Kimenyi (1980)
for detailed support of this claim.

   The second general means of forming non-patient passives is to derive passive
forms directly from the *n*-place predicate in such a way that the desired NP
is the subject. So where such derivational processes are well developed, as in
Austronesian, we find different morphological forms of passives of ditransitives
according to whether their subject is a patient or non-patient, and sometimes
we even find different morphologies on the verb depending on what sort of
non-patient is the subject. The basic pattern is illustrated below from Malagasy,
a subject-final language:

(59)   a. Nanasa ny lamba amin-ny savony Rasoa
            washed the clothes with-the soap   Rasoa
            'Rasoa washed the clothes with the soap'

        b. Nosasan-dRasoa amin-ny savony ny lamba
            washed-by.Rasoa with-the soap   the clothes
            'The clothes were washed with the soap by Rasoa'

        c. Nanasan-dRasoa   ny lamba ny savony
            washed.with-by.Rasoa the clothes the soap
            'The soap was washed the clothes with by Rasoa'

   Note in particular that the verb forms in (59b) and (59c) are different.
The form in (59b) tells us that the patient is the subject, and that in (59c)

tells us that some non-patient (finer distinctions can be made) is the subject. Philippine languages are particularly rich in the variety of verbal forms they present according to the semantic role of their subject. Example (60) below from Kalagan (Collins (1970)) is a relatively simple case; six different verbal forms are cited for Kapampangan in Mirikitani (1972) and up to twelve for Tagalog by Schachter and Otanes (1972), though not all verbs accept all forms.

(60)   a. K[um]amang aku   sa tubig na   lata adti balkon
           [ACTIVE]-get I(SUBJ) DO water with can on   porch
           'I'll get the water on the porch with the can'

       b. Kamang-in   ku   ya   tubig na   lata adti balkon
           get-PASS(PATIENT) I(AG) SUBJ water with can on   porch
           'The water will be got by me with the can on the porch'

       c. Pag-kamang   ku   ya   lata sa tubig adti balkon
           PASS(INSTR)-get I(AG) SUBJ can DO water on   porch
           'The can will be got water with by me on the porch'

       d. Kamang-an   ku   ya   balkon sa tubig na   lata
           get-PASS(LOC) I(AG) SUBJ porch   DO water with can
           'The porch will be got water on by me with a can'

4   Constructions that resemble passives

While many languages exhibit constructions that conform to the characteristics of passive constructions discussed above, other languages exhibit constructions that resemble passive constructions, and linguists are often not sure whether these constructions should be considered passives or not. In this section we discuss some examples of such constructions, and discuss briefly how they resemble passives and why they are generally not considered as such. In general, these constructions can be seen as lacking what we have taken to be the defining characteristic of passives: in a passive, the subject in the corresponding active is expressed by an element that is neither a subject nor an object in the corresponding passive or is not expressed at all; if it is not expressed, its existence is still entailed by the passive.

4.1   *Middles*

We have followed traditional practice in including the entailment of an agent as definitional of passives. Constructions which lack this characteristic but which otherwise resemble passives are generally called middles. The pair in (61) illustrate this contrast.

(61)   a. The ship was sunk
       b. The ship sank

While the passive in (61a) entails (61b), it has the additional entailment that there was some agent that caused the ship to sink, an entailment that is missing from the middle in (61b). This semantic difference coincides with the grammatical fact that, if the language allows expression of the agent in passive clauses, such is possible in a passive but not in a middle. Thus we can say *The ship was sunk by the enemy* but not **The ship sank by the enemy*.

While English does not employ any middle morphology, other languages do, and in some cases the morphology used for middles is similar to that used for passives. For example Classical Greek had a middle voice in addition to active and passive, and although the Greek middle has a range of functions that make it more than a middle in the technical sense used here, the middle and passive voices are identical in form in some tenses. Thus (62) can be interpreted either as a passive or as a middle.

(62)   Paú-omai
        stop-MIDDLE/PASSIVE.ISG
        I stop / I am stopped

The suffix *-ka* in Quechua sometimes has a middle interpretation, as in (63a), and sometimes a passive interpretation, as in (63b).

(63)   a. Punku kiča-ka-rqa-n
           door   open-MID-PAST-3
           'The door opened'

        b. Čuku apa-ka-rqa-n
           hat   take-PASS-PAST-3
           'The hat was taken'

The example in (63a) is a middle in that no agent is implied (though of course not excluded), while (63b) is passive since the meaning of the social action verb *take* does imply the existence of an agent. And in many languages, morphology that is basically reflexive is also used with both a middle and a passive function; the Spanish examples in (64) are analogous to the Quechua ones in (63).

(64)   a. Se   quemó   el dulce
           REFL burn.PAST.3SG the jam
           'The jam burned' (or 'The jam was burnt')

        b. Se   cumplieron   las promesas
           REFL fulfil.PAST.3PL the promises
           'The promises were fulfilled'

Nevertheless, despite the morphological and semantic similarities between middles and passives, the two can be distinguished by the diagnostic of whether an agent is entailed.

4.2   *Unspecified subject constructions*

A second and less common type of construction that is not always easy to distinguish from a passive is an unspecified subject construction, a construction with a subject whose meaning is roughly paraphrasable in English by 'someone' (or by 'someone or something'), or by 'they' or 'people' used generically. We use the term here to refer specifically to inflected forms of verbs, where there is a pronominal affix on the verb indicating that the subject is unspecified in this sense. For example, Kutenai has an unspecified subject suffix *-(n)am* that occurs on verbs as in (65).

(65)   a. qaky-am-ni   'taxa'
            say-UNSPEC.SUBJ-INDIC now
            'Someone said "Now!"'

       b. taxas ?at   qaky-am-ni   ?in   ?at
            then   HABIT say-UNSPEC.SUBJ-INDIC there HABIT

              n-uɫ   qanaɫunis-nam-ni
              INDIC-finish   travel-UNSPEC.SUBJ-INDIC
            'They say people used to travel that way'

This suffix in Kutenai is clearly an unspecified subject affix rather than a passive affix because it only occurs on intransitive verbs, never on transitive verbs; it cannot occur, for example, in a clause meaning 'someone killed him'. Its pronominal nature is brought out clearly by the fact that it also occurs on nouns to indicate an unspecified possessor, as in (66).

(66)   ?a·kitɫa?-nam
        house-UNSPEC.POSS
        'a house, someone's house'

There are various additional reasons, that we will not go into here, for saying that it is a pronominal affix rather than a voice affix.

   Now Kutenai also has a passive affix *-(i)ɫ* that is used with transitive verbs, illustrated above in (13), and in (67).

(67)   taxas paɫ   ?at   k-uniɫ-iɫ
        then   PTCL HABIT SUBORD-fear-PASS
        'Now people feared him'; 'he was feared'

How do we know that the suffix *-iɫ* in (67) is a passive suffix rather than an unspecified subject suffix? The answer to this becomes clear when we look at an example in which the semantic object is first or second person, as in (68).

(68)   hu   ṅ-iktuqu?-ɫ-ni
        1.SUBJ INDIC-wash-PASS-INDIC
        'I was washed'

If the suffix *-(i)ɬ* were an unspecified subject marker rather than a passive suffix, then (68) would literally mean 'someone washed me' and the first person argument would be treated as an object of the verb, with the first person singular object suffix *-nap* illustrated in (69a), but we get the first person subject proclitic *hu* illustrated in (69b).

(69)   a. ṅ-iktuqu-nap-ni
            INDIC-wash-1SG.OBJ-INDIC
            'He washed me'

        b. hu   ṅ-iktuquʔ-ni
            1.SUBJ INDIC-wash-INDIC
            'I washed him'

The fact that the first person argument behaves as a subject rather than as an object in (68) shows that the construction in (68) is a passive rather than an unspecified subject construction.

    Consider, in contrast, the Oneida construction in (70), which was mentioned earlier (see (9) above) as one construction that is used to fill the function served by a basic passive in languages without a basic passive.

(70)   úhkaʔ ok   waʔ-ukw-alahsʌtho-ʔ
        PRT   PRT   FACTUAL-UNSPEC.SUBJ:1.OBJ-kick-PUNCT
        'Someone kicked me'; 'I was kicked'

This Oneida sentence could be translated into English either by 'Someone kicked me' or by 'I was kicked', which mean approximately the same thing. The question is which of these two English translations more closely reflects the grammatical structure of the Oneida sentence: is it a transitive active sentence with an unspecified subject or is it a detransitivized passive sentence? In other words, does the morpheme *-ukw* indicate that it has an unspecified subject or does it indicate that the clause is passive? It turns out that a detailed analysis of Oneida morphology makes it clear that the former of these two possibilities is the correct answer: the morpheme *-ukw* is clearly a pronominal morpheme, with components that can independently be shown to indicate an unspecified subject and a first person object respectively.

    While the Oneida construction is fairly unambiguously an unspecified subject construction, there are constructions in other languages whose status is less clear. There is a construction in Algonquian languages whose analysis has been a source of debate for many years. This construction is illustrated by the example in (71) from Plains Cree.

(71)   ni-sa·kih-ikawi-n
        1-love-PASS/UNSPEC.SUBJ-SG
        'I am loved' *or* 'Someone loves me'

According to Wolfart (1973), the suffix *-ikawi* in (71) is an unspecified subject suffix, while according to Dahlstrom (1991), it is a passive suffix. The issue essentially revolves around the question of whether the first person singular prefix *ni-* is to be interpreted as a subject prefix, in which case the construction is passive, or an object prefix, in which case the construction is an unspecified subject construction. Analogous forms in the closely related language Ojibwa are interpreted by Bloomfield (1958) as passive forms, but Hockett, in the preface to Bloomfield's grammar, argues that this analysis is mistaken, that the forms in question are really unspecified subject forms. See also Dryer (1997b) for further discussion of the situation in Cree. A similar disagreement surrounds a prefix in Tlingit, which Story (1979) and Naish (1979) analyse as an unspecified subject prefix but which Boas (1917) analyses as a passive prefix. In fact, closely related languages can differ with respect to whether cognate constructions are passives or indefinite subject constructions. Mackay (1999) argues that the suffix *-kan* in (72) from Misantla Totonac is an indefinite subject suffix, since the notional object is represented by an object prefix.

(72)   kin-<u>iški</u>-kan-la
        1OBJ-hit-INDEF.SUBJ-PERF
        'Someone hit me'

But she notes that, in the closely related language Tepehua, the notional object is represented on the verb by subject affixes rather than object affixes with verbs bearing this suffix, arguing that it is a passive in Tepehua.

4.3   Inverses

A further type of construction which resembles passives is what are often called *inverses*. The prototype of this construction is again represented by Algonquian languages. The pair of Ojibwa examples in (73) represent what are traditionally referred to as 'direct' and 'inverse'.

(73)   a. aw   nini   w-gi:-wa:bm-a:n   niw   kwe:w-an
           that   man   3-PAST-see-DIRECT.ANIM.OBV   that.OBV   woman-OBV
           'The man saw the woman'

        b. aw   kwe: w-gi:-wa:bm-igo:n   niw   ninw-an
           that   woman   3-PAST-see-INVERSE.ANIM.OBV   that.OBV   man-OBV
           'The woman was seen by the man'

Using the more traditional terminology *actor* and *goal* to avoid begging the question as to what is the correct analysis, the direct and inverse in (73) differ as to which of the two elements has the grammatical status *proximate* (unmarked) and which is *obviative* (marked as OBV): in the direct clause in (73a), the actor

is proximate and the goal is obviative, while this is reversed in the inverse in (73b). There are interacting grammatical, semantic, and discourse factors governing the contrast of proximate and obviative, but, as a first approximation, we can say that the proximate element is the one that is more topical in the surrounding discourse. The crucial question here is whether the relation of direct and inverse should be considered an instance of active and passive. The answer to this question depends on whether the grammatical relations in direct and inverse are the same. If they are the same, in other words if the actor is subject in both clauses, then both are active and the inverse is not an instance of a passive. But if the goal is subject in the inverse, then this would appear to mean that the inverse is some sort of passive. Whether the goal is subject is something on which both positions have been taken, at least for different Algonquian languages; Dahlstrom (1991) defends the view for Cree that the actor is the subject in the inverse, thereby arguing against a passive analysis. Rhodes (1976) argues that in Ojibwa the goal is subject in the inverse, thereby arguing for a kind of passive analysis. However, even under the view that the goal is subject in the inverse, there is still a further question as to whether the construction should be considered a passive, revolving around the grammatical status of the actor and whether the clause is transitive. The transitivity of inverse clauses shows up most clearly in forms where at least one of the arguments is non-third person, illustrated by the Cree examples in (74).

(74)   a. ni-wa·pam-a·-w
             1-see-DIRECT-3
             'I see him'

         b. ni-wa·pam-ik-w
             1-see-INVERSE-3
             'He sees me'

In these cases, the direct is obligatory whenever the actor is higher than the goal on the person hierarchy 2nd > 1st > 3rd, as in (74a), and the inverse is obligatory whenever the actor is lower on this hierarchy, as in (74b). Under the view that the inverse is a passive, this means that passive is the sole way to express meanings in which a third person is acting on a non-third person, something that is unlike what we normally find among passives in other languages.

Our definition of passive requires that the agent, if expressed, be expressed neither as a subject nor as an object. This implies that passives will be intransitive, or, more accurately, that they will have valence one less than that of the corresponding active. But inverse clauses in Algonquian languages exhibit all the properties of being transitive: for example, the inverse clause in (74b) exhibits inflection for two arguments, just like the direct clause in (74a). In fact, Perlmutter and Rhodes (1988) propose that inverse clauses in Ojibwa involve

subject–object reversal, with the object becoming the subject and the subject becoming the object. If we define passives so as to require a decrease in valence, then the inverse clause in Ojibwa is not a passive, even if it involves a change in grammatical relations.

We can distinguish two sorts of inverse clauses in Algonquian languages, those in which both arguments are third person and they differ in terms of which is proximate, and those in which at least one of the arguments is non-third person. A number of other languages exhibit constructions that correspond to just one of these two types of inverses. Navajo (Athapaskan) exhibits an alternation without an apparent difference in transitivity:

(75)   a. łį́į́   dzaanééz yi-ztał
          horse mule   it.it-kicked
          'The horse kicked the mule'

       b. łį́į́   dzaanééz bi-ztał
          horse mule   it.it-kicked
          'The mule kicked the horse'

Unlike the Algonquian case, there is no grammatical distinction of proximate versus obviative in Navajo, although here the word order is crucial: the prefix *yi-* in (75a) indicates that the first noun is the agent, while the prefix *bi-* in (75b) indicates that the first noun is the patient. Other languages exhibit an inverse like the second sort of Algonquian inverse, in which at least one argument is non-third person. DeLancey (1981) cites the following example from Nocte (Tibeto-Burman).

(76)   a. nga-ma ate hetho-ang
          I-ERG   he   teach-1ST.3RD
          'I will teach him'

       b. ate-ma   nga-nang   hetho-h-ang
          he-ERG I-ACC   teach-INV-1ST.3RD
          'He will teach me'

In both examples in (76), the suffix *-ang* on the verb indicates that the verb has a first person argument and a third person argument. Without further indication as to which of these is subject and which is object, as in (76a), the subject is interpreted as the one that is higher on the person hierarchy, in this case the first person argument. If there is an inverse suffix on the verb, as in (76b), the subject is interpreted as the one that is lower on the person hierarchy, in this case the third person argument. Other languages that exhibit constructions that might be called inverses, of either of these two sorts, include Kutenai (Dryer (1994)), Nootka (Whistler (1985)), Cherokee (Scancarelli (1986)), and a number of other American Indian languages.

4.4   Antipassives

One further construction that resembles passive is antipassive, generally found in ergative languages. Like passive, antipassive involves a decrease in the valence of the clause, although in antipassives it is the patient-like constituent that is absent, or a non-argument. The pair of examples in (77) illustrate a pair of basic and antipassive clauses, respectively, from West Greenlandic.

(77)   a. arna-p   niqi-Ø   niri-vaa
           woman-ERG   meat-ABS   eat-INDIC.3SG.3SG
           'The woman ate the meat'

        b. arnaq-Ø   niqi-mik   niri-NNig-puq
           woman-ABS   meat-INSTR   eat-ANTIPASS-INDIC.3SG
           'The woman ate meat'

The fact that the antipassive in (77b) is intransitive is reflected by the fact that the verb cross-references only the absolutive *arnaq* 'woman', and this absolutive corresponds to the ergative, or transitive subject, in (77a). Intransitive subjects are absolutive.

   It should be noted that while antipassive constructions are typically associated with ergative languages, it is also not uncommon for ergative languages to have passives. For example, in addition to the antipassive construction in (77), West Greenlandic also has a passive construction. The examples in (78) illustrate an active sentence with its corresponding passive sentence.

(78)   a. inuit   nanuq   taku-aat
           people.ERG   bear.ABS   see-3PL.3SG.INDIC
           'The people saw the polar bear'

        b. nanuq   (inun-nit)   taku-niqar-puq
           bear.ABS   (people-ABL)   see-PASS-3SG.INDIC
           'The polar bear was seen (by the people)'

The differences between the active (78a) and the passive (78b) are: (i) the passive suffix on the verb in (78b); (ii) the optionality and ablative marking of *inunnit* 'people' in (78b); and the fact that the verb cross-references both noun phrases in the active (78a) but only *nanuq* 'bear' in the passive (78b).

5   The functional load of passive in grammars

We began this study by considering the functional role of passives in terms of foregrounding and backgrounding elements relative to actives. Passives certainly do have these functions, though they effect them in a rather specific way:

namely, by forming derived predicates whose argument structure differs in the ways we have considered from those they are derived from.

We have seen as well that languages vary considerably with regard to the productivity of their passives. Some languages have no passives at all; others present passives on a limited class of transitive and ditransitive verbs (those with patient objects), and not with intransitive verbs at all. Other languages on the other hand, such as many among the Bantu and Austronesian groups, essentially allow all verbs to passivize, and commonly a given verb will have several different passive forms according, for example, to the aspect of the derived structure or the semantic role of its derived subject. Given the productivity of passive in these languages, it would be surprising if passive formation did not interact with other rules of the grammar in regular ways.

One way to assess the importance of passive relative to other rules of grammar in a language is to ask what other syntactic/morphological operations can apply to passive structures, and why such operations effectively require that the structures they apply to be passive. As we mentioned earlier, passives, where they exist, are normally well integrated in the grammar, in the sense that major operations such as relative-clause formation, question-formation, and nominalizations typically can apply to passive structures. Some operations, however, such as imperative formation, often cannot. English is in this sense a typical language with passives.

In languages with highly productive passives, however, we find that the possibility of forming structures which are in principle independent of passives often depends on the existence of specific passives. Thus, if in English we lost the possibility of forming passives, we would not be obliged to change the way relative clauses are formed, or questions, or nominalizations, or reflexives; for, in fact, no major syntactic operation in English ever requires that the structure it operates on be passive. However, in Malagasy and the languages of the Philippines, by and large only main clause subjects can be relativized. That is, while we can literally say 'the man who washed the clothes', relativizing on the subject of 'washed', in Malagasy or Tagalog, we cannot literally say 'the clothes that the man washed' *(*ny lamba izay nanasa ny lehilahy)*. The expected structure can only mean 'the clothes which washed the man' – which doesn't make much sense. In any event, a relative clause is always understood in such a way that the head noun is the subject of the main verb of the subordinate clause. Thus, to refer to the clothes that John washed we must formally construe clothes as the subject of the subordinate clause. That is, we must say 'the clothes that were washed by John' using an appropriate passive form of 'wash' (cf. 59b, in which the patient is the subject). And to say 'the soap with which Mary washed the clothes' we must again use the passive form which presents the instrumental as the subject: literally 'the soap which was washed+with by Mary the clothes' (cf. 59c). Hence in Malagasy and the Philippine languages, if

we lost all passives, we would have to make very significant changes in the way relative clauses were formed. Similarly, the formation of constituent questions ('who did John kiss?') and 'clefts' ('it was Mary who John kissed') are not independent of passives in these languages. Thus, many major syntactic processes in effect require in certain contexts that what they operate on be passive. (See Keenan (1972) for a much more thorough discussion of this point.)

Similarly, in Bantu languages like Kinyarwanda we find that certain major processes such as relative-clause formation by and large (see Kimenyi (1980) for much discussion) operate only on subjects or direct objects. Thus to say 'the knife with which John killed the chicken' we must construe the subordinate clause as one in which 'knife' is either a subject or an object; it cannot be directly relativized as an oblique NP. So again, major syntactic operations depend on the existence of ways of forming derived objects and subjects in a way quite unlike English. Thus other constructions perform roles in grammar similar to those of passives.

It is often the case that if passives are highly productive in a particular language, then other syntactic processes may require the structures they operate on, in various cases, to be passive. Chapter 7 by Foley contains extensive further discussion on the functional load of passive constructions.

6   Suggestions for further reading

In addition to chapters 7 (by Foley) 3 (by Andrews), and 4 (by Dryer) in this volume, other relevant readings include Siewierska (1984), which discusses a wide array of issues surrounding passive constructions, and Klaiman (1991), which discusses passive constructions in relation to a number of related constructions, such as middles. Three anthologies which deal specifically with passives and related constructions are Shibatani (1988), Givón (1994), and Fox and Hopper (1994). There is much discussion of passives in the literature on Relational Grammar, including Perlmutter and Postal (1983), Postal (1986), and Blake (1990).

A number of chapters in *The World Atlas of Language Structures* (Haspelmath *et al.* (2005)) are relevant to topics discussed in this chapter, including Siewierska (2005b) and Polinsky (2005).

7 A typology of information packaging in the clause

William A. Foley

0   Introduction

All human languages possess syntactically variant ways to express what speakers intuitively feel is essentially the same conceptual event (see Grace (1987)). The syntactic variants available include word order differences, as in Yimas, a Papuan language of New Guinea (see chapter 3 by Andrews for an explanation of the symbols A, S, and P) –

(1)   kalakn klaki   ya-n-tay
        boy.SG parrot.PL 3PL.P-3SG.A-see
        'The boy saw the parrots'

– in which all possible permutations of this sentence are grammatical and have the same basic meaning as (1):

(2)   klaki ya-n-tay kalakn
        klaki kalakn ya-n-tay
        kalakn ya-n-tay klaki
        ya-n-tay klaki kalakn
        ya-n-tay kalakn klaki

Other possibilities are active–passive alternations exemplified by the English pair:

(3)   (a) The boy hit the ball
        (b) The ball was hit by the boy

Again, the two sentences contain the same basic information, 'boy hits ball', and one could say that the difference in syntactic structure seemingly has no effect. This is in marked contrast to some other shifts in word order in English: note, for example, that the sentences in (4) have exactly opposite meanings:

(4)   (a) Bill hit Fred
        (b) Fred hit Bill

Finally, there may be changes in the verb form of the sentence with consequent effects in the marking or order of associated NPs, but again with no difference in the basic meaning, as in these examples from the Mayan language Tzutujil:

(5) (a) jar aachi x-ɸ-uu-choy   chee^ʔ tza^ʔn machat
         the man   PAST-3SG.P-3SG.A-cut wood   with   machete

     (b) jar aachi machat x-ɸ-choy-*b'e*-ej   ja   chee^ʔ
         the man   machete PAST-3SG.S-cut-INSTR-SUFF the wood
         'The man cut the wood with the machete'   Dayley (1985)

Both the (a) and (b) sentences have the same basic meaning 'man uses machete to cut wood', but the syntactic form in which it is expressed varies: in the (a) sentence the verb is inflected transitively with prefixes for both the A argument (*uu-*) and the P argument (ɸ), and the instrument *machat* 'machete' appears in a prepositional phrase with *tza^ʔn* 'with' following the verb. In (b) the instrument appears as a bare NP preceding the verb, and the verb is inflected intransitively with only a null prefix for its sole S argument (ɸ), and carries the derivational suffix *-b'e* ('INSTR').

  In each of these examples, we see the same basic conceptual event 'packaged' in different ways through differences in word order, verbal morphology, case marking or grammatical function assignment. For example, the English active–passive alternation of (3) express the same basic information, the hitting of the ball by the boy, but the two NPs *the boy* and *the ball* have different kinds of 'salience'. In the first sentence the subject NP refers to the doer of the action and direct object NP refers to the entity affected by the action. In the second it is the NP denoting the affected participant which is the grammatical subject, while the NP denoting the doer occurs in a postverbal prepositional phrase. Note we normally understand the sentences in (3) as being about their subjects, so that we afford (3a, b) slightly different interpretations: (a) is saying something about *the boy*, while (b) comments about *the ball*. The same basic conceptual event is described in each sentence, but is expressed from a different perspective, 'packaged' in different ways.

  Why do languages possess variable ways of packaging the same event? Basically because, when people speak, they do so within a social context that includes previous speech, of themselves, of other interlocutors, even of long dead ancestors, in clichés, proverbs or even fables (Bakhtin (1981)). This context of speaking, together with wider sociocultural knowledge (Gumperz (1982)), is the background against which people construct and interpret the meaning and relevance of new utterances so that the utterances cohere to make up an intelligible discourse. A discourse is not merely a set of sentences

randomly strung together, but is rather a structured series, the development of which constitutes a coherent whole and is recognized as such by speakers of a language. Speakers therefore employ the various packaging options for clauses in the languages in order to ensure the coherence of the discourse. Each conceptual event described in the discourse will be presented in such a way as to foster the coherence of the discourse: thus, in a story about a boy and his afternoon play, we are much more likely to find a sentence like (3a) than (3b).

1   On verbal semantics and packaging options

1.1   *Conceptual events, participants and perspective*

We saw above a number of examples in which the same conceptual event has been presented in formally different ways, but the concept of conceptual event itself remains vague and undefined. Developing some ideas from Grace (1987), we may define a conceptual event as a basic level semantic description of an event regardless of any real-world, speech-time instantiation of it. Thus, the English verb *give* labels a conceptual event of someone causing an object to go into the possession of someone else, and this description applies to any instantiation of this label:

(6)   (a) Fred gave a book to Samantha
        (b) Fred is giving Samantha a book
        (c) Samantha has been given a book by Fred
        (d) Will a book be given to Samantha by Fred?

And this same basic description will apply to the Yimas verb 'give', although, of course, the label for the conceptual event will be different, *ŋa-* 'give' (PN = proper noun; Roman numerals in Yimas glosses indicate gender classes):

(7)   Yakayapan buk   na-n-ŋa-r-akn   Tamprak
        PN   book.V.SG   V.SG.P-3SG.A-give-PERF-3SG.DAT. PN
        'Yakayapan gave Tamprak a book'

Sometimes, the differences in expression of conceptual events may be greater than simply a change in the label. For example, English has a basic verb *kill* which describes a conceptual event of someone/something doing some action which results in someone else becoming dead: *Jane killed Alex*. Watam, another Papuan language of New Guinea, has no such verb; the conceptual event needs to expressed by a sequence of two verbs, one describing the causing action or event, which can be a general vague verb like *mo-* 'do' or a more specific one like *ruŋ-* 'hit', and another describing the result of dying:

(8)   Katewa Kas mo rugu-r   minik-rin
        PN   PN   ACC hit-DEP die-PAST
        'Katewa killed Kas'

That *ruŋ- minik-* ('hit die') 'kill' describes a unitary conceptual event in spite of its complex exponence is clear from that fact that (8) is a single clause; if we wanted two conceptual events with a cause–result relationship between them, a complex sentence with two clauses and an overt conjunction would be necessary:

(9)   Katewa Kas mo ruŋ-tape ma minik-rin
        PN   PN   ACC hit-CONJ 3SG die-PAST
        'Because Katewa hit Kas, she died'

   As examples like these from Watam make clear, the lexicalisation of conceptual events will vary quite widely from language to language (and even within a language: note *eat* versus *have a feed*). Ultimately, it is a function of the nature of a language's verbal lexicon, for a language's lexicalized verbs provide a systematic set of descriptions for states, events and actions, an inventory of basic conceptual events. Of course, the structures of verbal lexicons vary quite widely across the languages of the world, from many thousands of members for English, to around a hundred or so for Papuan languages like Yimas or Kalam (Pawley (1993)).
   Conceptual events, then, are those basic semantic descriptions of states, events or actions, be they of simple exponence, as in English, or complex, as in Papuan languages like Kalam, Watam or Yimas. A given event, then, can often be described in different ways across languages. Interestingly, even within a single language a similar phenomenon can occur; consider Fillmore's (1977) famous example of describing a commercial scene, say Sam's purchasing of a book from Jane. This can be described with different English verbs, each providing a different perspective on the event:

(10)   (a) Sam bought a book from Jane for $12
        (b) Jane sold a book to Sam for $12
        (c) Sam paid Jane $12 for a book
        (d) This book cost Sam $12

Buy presents the commercial event from the perspective of the buyer, while *sell* conversely presents it from that of the seller. *Pay* like *buy* takes the perspective of the buyer, but in addition attends to the money spent or the seller (while *buy* is more focussed on the goods). Finally, *cost* takes the perspective of both the goods and the money spent.

Perspective is a fundamental and pervasive fact of all human languages. English has many alternate pairs of lexical items and associated syntactic structures which present perspective differences.

(11)   (a) Clive likes Eric
        (b) Eric pleases Clive

(12)   (a) Bruce gave the ring to Sheila
        (b) Sheila took the ring from Bruce

(13)   (a) Malaria killed Alfred
        (b) Alfred died from malaria

(14)   (a) Alan lent the book to Sam
        (b) Sam borrowed the book from Alan

Note that there is a semantic difference between the (a) and (b) examples in each pair. It concerns the choice of the initiating and controlling participant in the event. This kind of difference is the essence of the perspective choices. Consider (14) which describes the transfer of a book from its permanent owner, Alan, to its temporary user, Sam. Note that in any typical act of the transfer of goods, the primary responsibility for the event can be ascribed to either of two participants, the source of the goods or their recipient. In (14a) Alan is considered by the speaker to be the primary performer and initiator of the action; hence *lend* is appropriate. Conversely, in (14b) Sam is described as the performer and initiator, and *borrow* provides that perspective. Similar distinctions obtain in (11–14). In all cases, the lexical alternatives present different perspective choices as to which participant is presented as the controlling, initiating or responsible performer of the event, what we will call the 'actor' of the event (comparable to Dowty's (1991) 'Proto-Agent'). Nor is this indication of perspective alternatives limited to lexical choices; in many languages the choice may be signalled morphologically:

(15)   Acehnese of Sumatra
        (a) lôn-bloe   buku nyan nibak jih
            1SG-buy   book that from   3SG
            'I bought a book from him'

        (b) ji-pu-bloe   buku nyan keu lôn
            3SG-CAUS-buy   book that to   1SG
            'He sold me a book'   Durie (1985)

Here the causative prefix *pu-* added to *bloe* 'buy' signals its converse 'sell'. *Bloe* takes the perspective of the buyer as actor, while the derived form *pu-bloe* takes that of the seller.

German shows a similar pattern with respect to the verbs 'buy'/'sell' and 'rent to/from'. In German there is a prefix *ver-* on the verb when the actor is the point of origin of the goods; the verb is unmarked when the actor is the recipient:

(16)   (a) Hans kaufte   drei Schildkröten von   Ursula
               PN   buy.PAST three turtles   from PN
              'Hans bought three turtles from Ursula'

         (b) Ursula *ver*-kaufte drei Schildkröten an Hans
               PN   sell.PAST three turtles   to PN
              'Ursula sold three turtles to Hans'

(17)   (a) Meine   Tochter mietete   die   Wohnung von
               my.NOM daughter rent.PAST DET.ACC apartment from
               dem   Arzt
               DET.DAT doctor
              'My daughter rented the apartment from the doctor'

         (b) Der   Arzt   *ver*-mietete die   Wohnung an
               DET.NOM doctor rent.PAST   DET.ACC apartment to
               meiner Tochter
               my.DAT daughter
              'The doctor rented the apartment to my daughter'

Note that the English verb *rent* has alternations much like those of German *mieten/vermieten* 'rent' except that there is no accompanying verbal morphology. The switch in preposition alone is sufficient to signal the change in perspective:

(18)   (a) Egbert rented the flat
         (b) Egbert rented the flat to Hortense
         (c) Hortense rented the flat from Egbert

The English verb *rent* is ambiguous with respect to the meanings of the German verbs *mieten/vermieten*. Note that, without a preposition, as in (18a), rent can mean either 'rent to' or 'rent from'; thus, Egbert in (18a) could either be the renter or the rentee. When a third argument is added in the form of a prepositional phrase as in (18b, c), the preposition chosen disambiguates the clause. Thus, if the preposition is *to*, the actor must be the point of origin of the transfer, as in (18b), while if the preposition is *from*, then the actor must be the recipient (18c). Thus, the choice of preposition in English performs the same function as the choice of *φ*- vs *ver*- in German (also with accompanying prepositions *von* 'from' / *an* 'to', as in (16) and (17)). English has relatively few examples of ambiguous verbs of the type exemplified by *rent*; in most cases the choice of

actor as goal vs point of origin is lexicalized through the selection of different verb stems, like *buy/sell*, *borrow/lend*, *take/give*.

The choice of actor is not the only alternative in perspective that most languages make available. Consider examples like the following:

(19)   (a) Howard robbed Frank of $100
         (b) Howard stole $100 from Frank

The English verbs *rob* and *steal* clearly present the same conceptual event – namely, because of Howard's theft, Frank is $100 poorer – but present it from different perspectives. With *rob* Frank as direct object is more directly involved in and affected by the action than with *steal*. Note that (20a) is fine, but (20b) is strange:

(20)   (a) Howard stole $50 from Frank, but Frank didn't know it
         (b) ?Howard robbed Frank of $50, but Frank didn't know it

The strangeness of (20b) is due to the fact that *rob* requires Frank to be affected by the theft; this is incompatible with denying his awareness of the event. The central affected participant of the event we will call the 'undergoer' of the event (comparable to Dowty's (1991) 'Proto-Patient').

Lexical alternatives like *rob/steal* are actually rather rare in English; normally perspective alternatives for undergoer selection are signalled by word order, verbal morphology or prepositional marking changes:

(21)   (a) The man loaded hay onto the truck
         (b) The man loaded the truck with hay

Note that the (b) example has a specific reading in which the truck is understood as being completely filled with hay.

(22)   (a) Harry sprayed paint on the wall
         (b) Harry sprayed the wall with paint

(23)   (a) John drained the water from the pool
         (b) John drained the pool of water

Note, again, that in the (b) examples the undergoer is normally interpreted as being completely affected, e.g. completely covered with paint or emptied of water, a reading which is not necessary in the (a) examples. In each of these, there is an NP which is understood as more completely affected when it is in direct object position than when it is the object of a preposition. Other examples of this type of alternation do not exhibit such marked differences:

(24)   (a) Sam hit the cane against the wall
         (b) Sam hit the wall with the cane

(25)   (a) Egbert gave the tortoise to Hortense
        (b) Egbert gave Hortense the tortoise

In (24) and (25) the NP in direct object position has greater discourse salience than the one that is the object of the preposition (we are thinking of it more than the other) but there is not a difference in how completely they are affected.

A canonical transitive verb like *break* will assign perspective to two NPs, actor perspective to one and undergoer perspective to another. The actor is the participant which performs, initiates or controls the event: it prototypically corresponds to the X participant in answer to the question 'what did X *do*?' Its most typical formal realization is as the subject of a transitive verb or the prepositionally or obliquely core-marked complement in the corresponding passive: *the man* (actor) *killed the duckling*; *the duckling was killed by the* man (actor). The single core argument of a good number of intransitive verbs (the unergative class) is also an actor: verbs like *laugh*, *say*, *speak*, *run*, *swim*, etc.; note that the single core arguments of such verbs do indeed answer the question 'what did X *do*?' 'X laughed, spoke', etc. The undergoer is the participant which is affected by the event: it prototypically corresponds to the Y in 'what *happened to* Y?' Its most typical formal realization is as the object of a transitive verb or the subject of the corresponding passive – *the man killed the duckling* (undergoer), *the duckling* (undergoer) *was killed by the man* – and, as we shall see, some verbs, labelled ditransitive, actually take two undergoers. So a canonical transitive verb like *kill* takes both an actor and an undergoer: *the man killed the duckling* – 'what did the man do?' (actor) 'the man killed the duckling'; 'what happened to the duckling?' (undergoer) 'the man killed the duckling'. There is also a class of intransitive verbs, the unaccusative verbs, whose single core argument functions as an undergoer rather than an actor: *fall*, *die*, *melt*, *get angry*, *fall pregnant*: 'what happened to Y?' 'Y fell, died, got angry', etc.

Let me represent the choice of actor and undergoer perspective as the assignment of features to participants: [+A] for the actor and [−A] for the undergoer. Universally it seems that a given verb root can only ever have one [+A] participant in its clause, but some verbs in some languages may permit multiple [−A] participants (as we shall see below, English is one such language). As we saw above, an intransitive verb can take only one NP marked for perspective, the choice being determined by semantics: a [+A] actor for unergative verbs like *swim* or *run*, and a [−A] undergoer for unaccusative verbs like *fall* or *happen*. There is a very small set of simultaneously unergative/unaccusative verbs like *roll* which take either a [+A] or [−A] participant, to be discussed further below. NPs which are specified either [+A] or [−A] by their governing verb are called *core* arguments; those which are not so specified are *oblique*. Oblique NPs must be governed by their own specific predicator, be it an adposition, case affix or

enclitic, or serial verb. Consider the following example:

(26)   Harry sprayed the wall with paint

The meaning of this sentence can be schematically represented in terms of the meaning of the conceptual event of the verb *spray* –

(27)   *spray*: someone propels a liquid through the air which causes the liquid to be on a surface

– which in the context of (26) means:

(28)   Harry propels paint through the air causing the paint to be on the wall

Since *spray* is lexically specified as a transitive verb in English, it takes both a [+A] and [−A] participant. The [+A] is straightforward, as the controlling and initiating participant *Harry* is the only candidate, but the [−A] is more complex. Both *paint* and *the wall* are potential candidates, as both are affected by the action of this event; the *paint* undergoes movement through the air and goes from the sprayer to *the wall*, while *the wall* goes from being unpainted to painted. Note both also answer the question 'What happened to X?': 'What happened to the paint?' 'We sprayed it on the wall'; 'What happened to the wall?' 'We sprayed paint on it'. Both participants have the semantic specification for [−A] undergoer status, but both cannot function as undergoers, i.e. in syntactic object function, because *spray* is lexically specified as a simple transitive verb, i.e. it allows a single syntactic object or [−A] undergoer. This is quite unlike true ditransitive verbs like English *give* which do allow multiple objects or [−A] undergoers. In (26) it is *the wall* which has been chosen as undergoer. This forces *paint* to be oblique and thus governed by the preposition *with*, which has a lexical meaning like 'something that someone uses/moves'. Note that the meaning of *with* is compatible with the overall meaning of *spray*, so that (26) is a coherent clause. If a preposition with a radically different meaning were selected, like *at*, a clash of meanings between it and *spray* might result, yielding an incoherent clause:

(29)   ?Harry sprayed the wall at paint

1.2   *Parameters governing actor choices*

The actor or [+A] participant is prototypically the causing, controlling, initiator of the action. In most cases the identification of the [+A] NP is clear, but not always.

Sometimes there are two potential causers of the action as in (30):

(30)   (a) John opened the safe with the key
        (b) *The key opened the safe with John

Thus, while both *John* and *the key* are causing, performing participants of the actions, i.e. John causes the key to cause the safe to become open, only John is eligible to be [+A], *the key* being consigned to oblique status. Note that if John is absent in the sentence, *the key* can easily assume the [+A] function:

(31)   The key opened the safe

This demonstrates the need for an explicit statement of accessibility of participant types for the [+A] function, with some types outranking others (see Foley and Van Valin (1984), and also Dowty (1991)). The strongest parameter for [+A] status, outranking all others, is volitional causation or initiation of the action. This accounts for why *John* outranks *the key* in (31). In some languages like Acehnese (Durie (1985)), a participant can normally only function as [+A] when it is a volitional performer. In Acehnese the [+A] NP is marked by a verbal proclitic:

(32)   (a) Si   Ali *ji*-timbak   si   Mat
            title NP 3SG[+A]-shoot title PN
            'Ali shot Mat (deliberately)'

        (b) asee nyan ka   *ji*-poh
            dog   that PERF   3SG[+A]-beat
            'He beat that dog'

Many languages are less restrictive than Acehnese and allow any causing or initiating person, instrument or force to be assigned [+A], although, of course, they remain outranked by controlling volitional causers, as in (30). Both English and Yimas are like this:

(33)   English
        (a) The key opened the door
        (b) An earthquake levelled the city
        (c) Malaria killed John
        (d) Smoke blackened the roof

(34)   Yimas (the Roman numerals indicate noun class assignment)
        ikn   antki   ya-n-tal-urkpwica-t
        smoke.V.SG thatch.IV.PL IV.PL[−A]-V.SG[+A]-CAUS-blacken-PERF
        'Smoke blackened the roof'

Of course, these are weaker choices for [+A], so, not surprisingly, alternatives in which they are not [+A], but oblique NPs, are possible:

(35)   (a) John died from malaria [compare (33c)]

        (b) ikn-an   antki   ya-urkpwica-t [compare (34)]
            smoke.V.SG-OBLIQ thatch.IV.PL IV.PL[−A]-blacken-PERF
            'The roof got blackened from the smoke'

Now consider a situation in which Patrick has a phobia of snakes, which could be described either as (36a) or (36b):

(36)   (a) Patrick fears snakes
        (b) Snakes scare Patrick

Note that the identification of *snakes* as [+A] in (36b) is straightforward – 'What do snakes do to Patrick?' 'They scare the hell out of him' – but this is not the case with (36a): 'What does Patrick do?' *'He fears snakes'. Another difference between the two sentences is that *fear* is a stative verb while *scare* is dynamic, hence the former fails to occur in the progressive or imperative: *Patrick is fearing snakes*, *Fear snakes!*; *The snakes are scaring Patrick, scare the snakes!* These data demonstrate that while *snakes* in (36b) are indeed the responsible causers and controllers of the event – in a word, [+A] participants – this is not the case with Patrick in (36a). Thus, while *Patrick* has the same formal syntactic properties as *snakes*, i.e. those of subjects of formally transitive verbs, it is not a true or prototypical [+A] participant. This claim is further buttressed by the fact that it cannot function as the object of the preposition *by* in a corresponding passive, a diagnostic property of all [+A] participants of transitive or ditransitive verbs: *snakes are feared by Patrick*.

Patrick in (36a) illustrates the semantic role of experiencer, a sentient participant having a sensory experience of a perceptual, cognitive, emotional or bodily event or state. While experiencers are commonly realized like [+A] participants, i.e. as subjects of transitive verbs, they are not true prototypical [+A] participants because they are not responsible causers of the state or event. Many languages exhibit this wide-spread syntactic conflation of true [+A] participants and experiencers, across a wide range of types of experiential states:

(37)   English
        (a) I saw the accident
        (b) Fran knows Sam very well
        (c) Egbert loves Ryan
        (d) I feel sad

(38)   Yimas
        (a) impa-ka-tay   ŋaykumprum
            3DU.P-1SG.A-see   woman.II.DU
            'I saw the two women'

        (b) wapun   ama-na-pay-n
            happiness.V.SG   1SG.S-PROG-carry-PRES
            'I'm feeling happy'

Strong evidence that these experiencers are not true [+A] participants comes from that fact that there exist paraphrases of such sentences as (37) and

(38) in which the experiencer appears in a non-subject function without a passive derivation, i.e. the experiencers are realized as obliques or syntactic objects:

(39)   (a) The accident was visible to me (from my window) [compare (37a)]

        (b) Sam is well known to Fran [compare (37b)]

        (c) Yimas
            wapun   na-ŋa-na-t-n [compare (38b)]
            happiness.V.SG V.SG.A-1SG.P-PROG-do/feel-PRES
            'I'm feeling happy'
            (literally 'happiness does me')

This is in marked contrast to true prototypical [+A] participants. Whenever they occur, they must function as the syntactic subject, if no passive derivation has applied; no other syntactic function is available to them: *The mouse was eaten at/to/from the cat*; *The city was levelled at/to/from the earthquake*. It is also worth pointing out that there are many languages, for example the Daghestanian languages of the Caucasus, which never allow experiencers to take on the same syntactic properties as prototypical [+A] participants. In such languages, they typically occur in an oblique case such as locative or dative, while true [+A] participants take core cases like ergative or nominative. This again demonstrates that experiencers of sensory states are not true [+A] participants.

    A final parameter governing choices for [+A] is movement. Any participant which moves relative to another participant is eligible to be [+A], even if the participant is inanimate and hence totally incapable of volition and movement under its own will. English illustrates this well:

(40)   (a) The train passed the village
        (b) Water reached the bridge

This parameter seems to be universal; it appears as if a kind of metaphorical volition is ascribed to moving objects. Even in Acehnese, which otherwise has a strong restriction of [+A] to volitional causing participants, moving objects are marked with the [+A] proclitics (Durie (1985)):

(41)   (a) apolô ji-pho   u buleuen
            Apollo 3SG[+A]-flew to moon
            'Apollo flew to the moon'

        (b) ie   naya pi ji-teuka
            water big   too 3SG[+A]-arrive
            'The floods came'

1.3   Parameters governing undergoer choices

In the same way that the agent of an action outranks all others for [+A], participants which undergo a change in state outrank all other contenders for [−A]. Note the behaviour of a prototypical change-of-state verb like *break*:

(42)   (a) John broke the vase with the cane
          (b) John broke the cane against the vase

Unlike earlier paired examples (21–5) involving perspective alternations for undergoer, (42a) and (42b) are not at all synonymous. In (42a) it is the vase which undergoes a change in state, i.e. breaks, while in (42b) it is the cane. This demonstrates that verbs like *break*, which entail a change in state of a participant, i.e. something comes to be in a resulting state – broken, frozen, dead, wide, etc. – do not allow alternations for undergoer. The participant undergoing the change in state must be the [−A], and this is a pattern which appears universal.

Other participants which may be [−A] are those which undergo a displacement in space, a change in location rather than state. Again, these are universally available to [−A]:

(43)   (a) John hit the cane against the wall
          (b) John drained the water from the pool
          (c) John sprayed paint on the wall
          (d) John loaded hay onto the truck
          (e) Yimas
              namarawt   awtmayɲ-ɲan   na-ŋa-tpul
              person.1.SG sugar.cane-OBLIQ 3SG[+A]-1SG[−A]-hit
              'The man hit me with (stalks of) sugar cane'
          (f) Acehnese
              lôn-rhöm   batèe   bak boh mamplam
              1SG[+A]-throw stone[−A] at fruit mango
              'I threw a stone at a mango'   Durie (1985)

However, these are less prototypical [−A] participants, so that many languages, including English, but not Yimas, allow alternatives of perspective in sentences like these. This involves the next parameter for availability for [−A] assignment, affectedness. If a participant can be viewed as being in some way causally affected by the action of another participant in the clause, typically the [+A] participant, it can assume the [−A] function. Note that entities which undergo a change in location such as the undergoers in (43) can clearly be claimed to be affected by this spatial displacement, and so are [−A] by this affectedness parameter, but it is also possible for the locations in these displacements to be affected. All of the English examples in (43) have such alternatives:

(44)   (a) John hit the wall with the cane
        (b) John drained the pool of the water
        (c) John sprayed the wall with paint
        (d) John loaded the truck with hay

Again, it is the syntactic object which is [−A], so while in (43a) *the cane* is [−A], in (44a) it is *the wall* which is [−A]. The [−A] NPs in (44) usually entail a completely affected meaning as a result of the action, as opposed to their interpretations as obliques in (43). It is important to note that, while both the object displaced and the location in the examples of (43) and (44) are potential undergoers by virtue of the affectedness parameter, and both answer the question 'what happened to X?' ('What happened to the cane?' 'John hit it against the wall', 'John hit the wall with it'; 'What happened to the wall'? 'John hit the cane against it', 'John hit it with the cane'), in any given sentence *only one* can be the undergoer, because the verbs involved are simple transitive verbs which can only occur with a single syntactic object, the [−A] participant. Because of this *syntactic* constraint, the other possible [−A] must appear in an oblique function, in spite of the fact that *semantically* it meets the requirements for undergoerhood.

   Cross-linguistically, however, there are verbs that allow multiple choices of [−A] NPs. These are ditransitive verbs like *show*, *tell*, *offer*, of which undoubtedly the most prototypical is *give*. *Give* and other true ditransitive verbs in English exhibit an alternation not yet discussed:

(45)   (a) Egbert gave the snake to Mildred
        (b) Egbert gave Mildred the snake

Give is a transfer verb like *load* or *drain*, but, unlike these, the object which is transferred not only undergoes a change in location, but prototypically a change in ownership as well. In (45a) *give* behaves like a standard transitive verb, with [−A] assignment to the object which undergoes change in ownership. The new owner appears as an oblique, in a prepositional phrase with *to*. The new owner is obviously affected by the transfer by coming into possession of the object given, is therefore a potential [−A], but because *give* is a simple transitive verb in this clause frame and the [−A] function is already taken by the object given, it must occur in oblique function. Example (45b) illustrates an as yet unseen pattern, in which the new owner now functions as the [−A]. Of course, the object given still remains affected by the transfer, but in these cases, as opposed to the earlier examples of (43) and (44), it is not realized as an oblique but remains an unmarked core argument of the verb, and hence must be assigned perspective by the verb, in this case [−A]. We can define ditransitive verbs as that class which is lexically specified as being able to take two [−A] participants, i.e. two syntactic objects, rather than the usual one of transitive verbs. English is by no

means the only language to illustrate the alternation in (45); it is exemplified as well by Nengone of New Caledonia (Tryon (1967)) and Lango of Uganda (Noonan and Bavin-Woock (1978)), among many other languages:

(46)   Nengone
         (a) inu či   kanon ɔre tusi   du bɔn
              I   PRES   give   the book to 3SG
              'I give the book to him'

         (b) inu či   kanon bɔn re   tusi
              I   PRES   give   3SG the book
              'I give him the book'

(47)   Lango
         (a) lócà òmí᷈ɔ mɔ̀t bɔ̀ àtɪ́n
              man gave   gift to child
              'The man gave a gift to the child'

         (b) lócà òmí᷈ɔ àtɪ́n mɔ̀t
              man gave   child gift
              'The man gave the child a gift'

In all three languages, English, Nengone and Lango, *give* occurs in two clause structures: in the first, illustrated by the (a) examples, *give* is a simple transitive verb with a single [−A] participant – the thing given, realized as the syntactic object – and the recipient appears as an oblique constituent, hence out of the perspective imposed by the verb. In the second structure, *give* is a ditransitive verb, with two [−A] participants, the gift and the recipient, functioning as syntactic objects, as in the (b) examples. Both participants appear with the diagnostic grammatical properties of core NPs immediately after the verb and hence must be with the perspective assignment of the verb. Neither takes an oblique marker such as an adposition or case ending. It is important to remember that the notion of [−A] has both semantic and syntactic diagnostic properties. A [−A] NP must answer the question 'What happened to X?' *and* function in a core syntactic position, prototypically a syntactic object of a transitive or ditransitive verb or the subject of a corresponding passive. The semantic property on its own is not sufficient, for example the recipient of an act of transfer such as is entailed by a verb like *give* (example (45)) is always affected by the action: they come into possession of an object and hence meet the semantic requirements for [−A] status. It may, however, not achieve that status if *give* is functioning as a simple transitive verb (45a) and its single [−A] slot is already filled by the object given. In such cases it will necessarily appear in oblique function. If, however, *give* is ditransitive with two [−A] slots, this strictly syntactic constraint will no longer block the recipient from assuming [−A] status. Both the semantic and

the syntactic constraints for [−A] status then can be satisfied and the recipient will appear as such.

It is important to note that *give* is not potentially ditransitive in every language; in fact there may indeed be languages without ditransitive verbs at all. For example, in Tolai, an Austronesian language of New Britain (Mosel (1984)), *tar* 'give' is transitive, with the gift as the [−A] participant; the recipient must be specified as an oblique in a prepositional phrase:

(48)   i   ga   tar ia tai tura-na
           3SG REM give 3SG to brother-3SG.POSS
           'He gave it to his brother'

Many languages, notably the classical Indo-European languages but a number of others as well, have a distinct oblique case form, the dative, which marks the recipient with *give*. The gift typically takes the diagnostic core grammatical properties, the syntactic and morphological properties of the [−A] participant, such as grammatical case (see chapter 3 by Andrews) like accusative or absolutive or syntactic properties like objecthood, while the dative NP behaves like an oblique for most purposes:

(49)
(a) German
   die   Frau   hat den   Kindern   das   Buch gegeben
   NOM.SG woman has DAT.PL children-DAT.PL ACC.SG book given
   'The woman gave the book to the children'

(b) Ingush
   da·s   woɛaa kita·b   dɛa-lu
   father.ERG son.DAT book.ABS away-give
   'Father gives son a book'   Nichols (1994)

Some of these languages may have a small class of ditransitive verbs not including *give* (for example, like German *lehren* 'teach' with two accusative, i.e. [−A], NPs), or they may lack the class entirely, like Tolai. This is one area of typological studies which is greatly in need of further intensive work. It is worth noting that most languages lacking ditransitive verbs – i.e. having no verbs that can take two [−A]s, two NPs in object function – always present the recipient in the oblique frame. It would be worth knowing if there are many languages possessing only transitive *give* which do so in the other possible frame, i.e. something like *Egbert gave Mildred with the snake*, in which the recipient is [−A] and the gift is oblique. Are there many languages like this? Straits Salish (Jelinek and Demers (1994)) seems to be one such language:

(50)

'oŋəs-t-oŋəɬ=sxʷ   'ə   cə   kʷən-t-əxʷ

give-TRANS-1PL[−A]=2SG[+A] OBLIQ DET take-TRANS-2SG[+A]SUBORD

'You gave us (with what you caught)'

But such languages are decidedly rare and the question that needs to be answered is: why? (Chamorro (Cooreman (1988)) may be another language which illustrates this pattern, but the facts are complex.)

At the opposite extreme from languages like Tolai are those in which *give* and related verbs are necessarily ditransitive, with no possible transitive equivalents. Bantu languages around Lake Victoria in Africa are commonly like this, for example KiHaya of Tanzania (Hyman and Duranti (1982)) (phonological rules drop final vowels preceding words beginning in a vowel):

(51)   (a) a-ka-h'   ómwáán' ébitooke
            3SG.SUBJ-PAST-give child   bananas
            'He gave the child bananas'

       (b) a-ka-siig'   ómwáán' ámajûta
            3SG.SUBJ-PAST-smear child   oil
            'He smeared the child with oil'

Note that both postverbal NPs appear as bare NPs with neither indicated as oblique. This suggests that both are core NPs and function as [−A]. Further syntactic and morphological evidence confirms this: the two NPs are equivalent in every grammatical respect; they can occur in either order following the verb (51a and 52); they can both be subjects of corresponding passive sentences (53a, b); and both can be replaced by [−A] verbal pronominal agreement prefixes (54a, b, c). No grammatical property can be found which distinguishes between them, and they must both be identified as undergoers

(52)   a-ka-h'   ébitook' ómwáana
         3SG.SUBJ-PAST-give bananas child
         'He gave the child bananas'

(53)   (a) omwáán' a-ka-háá-bw'   ebitooke
            child   3SG.SUBJ-PAST-give-PASS bananas
            'The child was given bananas'

       (b) ebitooke bí-ka-háá-bw'   ómwáana
            bananas   3PL.SUBJ-PAST-give-PASS child
            'The bananas were given to the child'

(54) (a) a-ka-mú-h'   ebitooke
          3SG.SUBJ-PAST-3.SG[−A]-give bananas
          'He gave him bananas'

      (b) a-ka-bí-h'   ómwáana
          3SG.SUBJ-PAST-3PL[−A]-give child
          'He gave them to the child'

      (c) a-ka-bi-mú-ha
          3SG.SUBJ-PAST-3PL[−A]-3SG[−A]-give
          'He gave him them'

Many other languages with ditransitive verbs, perhaps most, fail to mark both [−A] participants with pronominal agreement affixes. In such situations it is invariably the recipient which is marked by verbal agreement, while the gift remains core, i.e. does not have the morphology or syntax of obliques such as adpositions or oblique cases, even though it fails to register on the verb through agreement. Many Bantu languages are like this, in contrast to KiHaya:

(55) Chi-Mwi·ni
      (a) ni-m-pełe   Ja·ma kujá
          1SG.SUBJ-3SG.[−A]-give PN   food
          'I gave Jama food'

      (b) *ni-'i-pełe   Ja·ma kujá
          1SG.SUBJ-3SG[−A]-give PN   food
                                        Kisseberth and Abasheikh (1977)

Example (55a) in which verbal agreement is with *Ja·ma*, the recipient, is gram-matical, but (55b) is not, because agreement is with the object transferred (the contrast in prefixes *m-* and '*i-* represents a difference in gender for the two nouns). Note that *kujá* 'food', however, still remains core; attempting to mark it oblique results in ungrammaticality:

(56) *ni-m-pełe   Ja·ma ka· kujá
      1SG.SUBJ-3SG[−A]-give PN   with food

Many languages around the world exemplify this pattern, which we might term *restricted ditransitivity*.

Just as we noted earlier that there seem to be very few languages with only transitive *give* that allow the recipient to be [−A], i.e. *John gave Mildred with the snake*, there also seems to be a lacuna of languages with restricted ditran-sitivity in which the object transferred is cross-referenced as [−A], but the recipient is not. There seems to be a converse, mirror image pattern here. In exclusively transitive languages, only the object transferred is eligible to be the

[−A] participant; while in restricted ditransitive languages, only the recipient is eligible for verbal agreement as [−A], although the object transferred otherwise remains a core [−A] participant. Again, are there languages which have the equivalent of (57)?

(57)   man books child 3SG[+A]-3PL[−A]-give
          'The man gave the child books'

And, again, if not, why not? One possible explanation is the higher animacy of the recipient compared to the object transferred, but what is the normal cross-linguistic pattern in restricted ditransitive languages for cases when both are equal in animacy: *the man gave the orphan to the parents*; *I gave him to them*? Clearly, there is a great deal we do not yet know about ditransitivity and perspective alternatives for [−A].

1.4   *Intransitive verbs and the unaccusative/unergative split*

Intransitive verbs, those that subcategorize for a single core argument, can present especially interesting complexities when it comes to the choice of [+A] or [−A] perspective. Potentially, the choice of either [+A] or [−A] is available, so how is the selection decided? Essentially, intransitive verbs divide into two broad classes: what are termed unergative verbs, which denote activities or actions for which the single argument is the causer or initiator, i.e. X *does* something, such as *swim*, *walk*, *run*, *ascend*, *cry*, *chat*, and unaccusative verbs, which denote states or processes that the single argument of the verb is in or undergoes, i.e. something *happens to* Y, such as *fall*, *melt*, *break*, *slip*, *be thirsty*, *be sleepy*. The prototypical argument of an unergative verb is an agent or causer, so they take a [+A] participant, while that of an unaccusative verb is affected by a state or change in state, so that they occur with a [−A] participant. In English, this contrast shows no overt morphosyntactic difference (though there are more subtle, covert syntactic differences, see Bresnan (1994)); so a pronominal argument of verbs from either of these classes has the same form: *I swam*, *I slipped*. However, many languages of the world do have overt grammatical differences for these two classes. As with transitive verbs, the parameters which determine the choice of [±A] can vary across and within languages, so let's look in greater detail at the types of variation found.

Acehnese, as we have already seen, has strong restrictions on the selection of [+A] NPs for transitive verbs, namely the argument must be a volitional controlling initiator of the action, an agent. Exactly the same constraint applies to the choice of [+A] for intransitive verbs: only performers of actions that can be construed as volitional can be [+A] participants, which are realized with

the agentive proclitics. If the event is not volitionally accomplished, it is taken as happening to someone, to be an unaccusative verb with an associated [−A] participant, which is expressed as a pronominal enclitic. An example of each type follows:

(58)   unergative verb: [+A]   unaccusative verb: [−A]
         geu-jak   rhët-geuh
         3SG[+A]-walk   fall-3SG[−A]
         'He walks'   'He falls'   Durie (1987)

Other examples of unergative and unaccusative verbs are:

(59)   *unergative verbs*   *unaccusative verbs*
         ék   'ascend'   *rô*   'spill'
         döng   'stand'   *beureutôh* 'explode'
         klik   'cry'   *trôh*   'happen, arrive'
         hah   'open mouth'   *beukah*   'be broken'
         marit   'talk'   *habêh*   'be finished'
         kira   'think'   *êk*   'like'
                                     caröng   'be clever'
                                     gli   'be ticklish'   Durie (1985)

Acehnese focusses on prototypical [+A] parameters like volition and control and uses that as the basis of the split between unergative and unaccusative verbs. Other languages with a similar basis for the intransitive verb split are Caddo and Lakhota (both Mithun (1991)) and Tsova Tush (Holisky (1987)). It is also just as possible to select typical [−A] parameters like change of state or affectedness as the basis of the split. Languages which make use of this parameter will contrast verbs which denote changes of state, or states with a significant or long-term (unaccusative) effect, from those which do not (unergative). Tolai of New Guinea (Mosel (1984)) reflects an unergative–unaccusative split along these lines:

(60)   (a) unergative [+A]
             a   tutana i   vana
             ART   man   3SG go
               [+A]
             'The man went'

         (b) unaccusative [−A]
             i   ga   kubur a   lama
             3SG REM grow ART coconut
                                    [−A]
             'The coconut grew'

Tolai word order rules specify that [+A] participants precede the verb, while [−A]s follow, as (60) demonstrates. The same rule applies to transitive verbs (see example (48)) with word order [+A] – V – [−A]). Other examples of unergative and unaccusative verbs in Tolai include:

(61)   *unergative*   *unaccusative*
         momo 'drink'   *io*   'burn'
         ruk   'enter'   *por*   'be finished'
         rovoi 'hunt'   *dudu*   'sink'
         peke   'excrete'   *kapa*   'be clear'
         kikita 'beat'   *papala* 'be open'

   This contrast between unergative and unaccusative as revolving around the notion of change of state rather than agentive control is closely correlated with a common effect of aspect on the semantic bases of the split. This is perhaps not too surprising; the close correlation between perfect or perfective aspect and resultant changes of state is well known (Comrie (1977)), as is the converse relation between imperfective aspect and activities and actions (DeLancey (1981)). An agent does an activity or carries out an action, the duration of which can be highlighted by imperfective aspect. A perfective or perfect aspect, on the other hand, indicates the completion of an event process and highlights how this event or process has brought about a resultant change of state on the affected participant. Therefore, we should not be surprised to find effects of aspect contrasts skewing the basis of the unergative–unaccusative split. This is the basis of the much discussed unergative–unaccusative contrast in Italian (Van Valin (1990)):

(62)   (a) imperfective: unergative with auxiliary *avere* 'have'
              Luisa ha   corso nel   parco per un' ora
              PN   has run   in the park for an   hour
              [+A]
              'Luisa ran in the park for an hour'

         (b) perfective: unaccusative with auxiliary *essere* 'be'
              Luisa è   corsa a   casa   in un' ora
              PN   be run   to   house in an   hour
              [−A]
              'Luisa ran home in an hour'

The unergative form with *avere* in (62a) is imperfective, highlighting the duration of the event over a span of an hour; no completed end point is asserted. Conversely, the unaccusative with *essere* in (62b) is perfective; Luisa has arrived home after running there for an hour. Note that in these Italian examples the

semantic basis of the unergative–unaccusative split is extended well beyond its normal correlation between the contrast between [+A] and [−A] respectively. By the usual semantic definitions, 'What did X do?' 'What happened to X?', Luisa in both (62a) and (b) is a [+A] participant, because both sentences can answer the question 'What did Luisa do?' The contrast that seems to be expressed here is akin to a scale of prototypical, and perhaps most active, [+A] status. The more actively involved [+A] participant – i.e. the one still involved in the action in which he is the controlling causer, so the event is described in imperfective aspect – exhibits the diagnostic properties of the unergative construction, the one typically linked to [+A] participants; while a less involved [+A] participant – one which could be viewed to no longer be in control of the event because it is now finished, i.e. in perfective aspect – occurs in an unaccusative construction, the one prototypically linked to [−A] participants. While such unaccusative [+A] participants are not [−A]s, they are clearly seen as less potent [+A] participants than unergative ones, and the language utilizes the unergative–unaccusative contrast to signal this subtle difference between types of [+A] participants.

2   On argument structure and pivots

2.1   *The nature of argument structure*

Following Manning (1996), we treat argument structure as a strictly *syntactic* representation of the arguments of the verb, in contrast to the lexical *semantic* representations discussed in the previous section, with their variants of perspective. Argument structure simply lists the number of arguments that a verb might have:

> single argument, monovalent verbs: *swim* <x>; *fall* <x>
> two arguments, bivalent verbs: *break* <x, y>; *kill* <x, y>
> three arguments, trivalent verbs: *spray* <x, y, z>; *give* <x, y, z>

There seem to be no basic verb roots in any language with more than three arguments. (Verbs of commercial transaction like *buy* are often thought of as verbs with more than three arguments – *Mary bought a boa constrictor from Sue for $500* – but this is clearly not the case, for it is never possible for more than three NPs to function as core in a given clause: *Mary bought Egbert a boa constrictor*. In fact, verbs like *buy* are simple transitive verbs which can become ditransitive through a process of benefactive argument addition, as above. Without this derivation, *buy* is simply *transitive – Mary bought a boa constrictor* – all other NPs are necessarily oblique and not subcategorized by the verb. It may be a *semantic* fact that an act of buying requires money and a place/person from whom an object is bought, but these are real-world facts

about such events, not *syntactic* constraints on how many arguments are linked to the verb in its lexical entry.) There are basic precedence relations among the arguments in an argument structure which are determined by the linking to their lexical semantic representation and [A] feature assignments. The algorithm that links these is simple:

$$[+\text{A}] > [-\text{A}] > [\ ],$$

i.e. core arguments outrank obliques and, among core arguments, the [+A] participant is foremost. So, for a sentence like

(63)   John sprayed the wall with paint
         [+A]   [−A]   [ ]

we have the following linking

      lexical semantic structure: *spray*
      John cause paint to move to be at the wall
       | | |
    [+A]   [ ]   [−A]

      argument structure: *spray* <x, y, z>
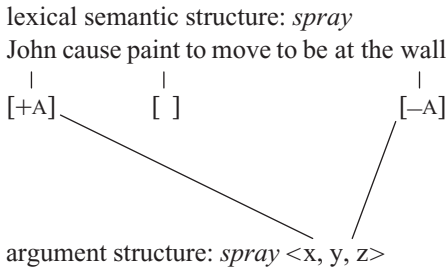
whereas for

(64)   John sprayed paint on the wall

we find the following linking pattern:

      lexical semantic structure: *spray*
      John causes paint to move to be at the wall
       | | |
    [+A]   [−A]   [ ]

      argument structure: *spray* <x, y, z>
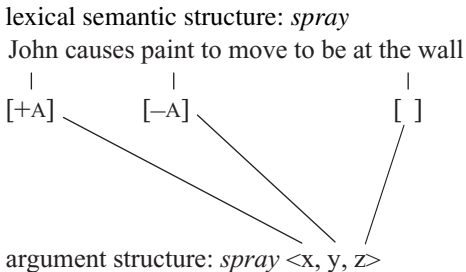
As we can see, the nature of argument structure and the precedence relations among the arguments is normally very simple to determine via the linking algorithm of $[+\text{A}] > [-\text{A}] > [\ ]$. The only area of potential difficulty is that of true ditransitive verbs which assign two $[-\text{A}]$ features. What are the precedence relations, if any, between these two? Basically there are two possibilities:

the two NPs are treated identically (symmetrical languages) or they are treated differently (asymmetrical languages), with the recipient participant always out-ranking the object transferred. Bantu languages exhibit significant variation in this area (see Alsina (1993) for a full discussion of the range of possibilities for this parameter in Bantu languages). Differences of animacy between the two [−A] NPs often have important effects on whether they behave symmet-rically or asymmetrically: when both NPs denote animate or, even more so, human beings, they are highly likely to demonstrate symmetrical behaviour, as in Shona (Hawkinson and Hyman (1974)). KiHaya, for example, is a com-pletely symmetrical language as examples (52–4) demonstrate: both [−A] NPs of *ha-* 'give' behave identically with regard to a number of morphological and syntactic properties. Chi-Mwi·ni (examples (55–6)), on the other hand, is an asymmetrical language – only the recipient of ditransitive 'give' is accessible to many grammatical possibilities, such as (Kisseberth and Abasheikh (1977)):

(65)   verb agreement (same as 55)

        (a) ni-m-peɫe   Ja·ma kujá
             1SG.SUBJ-3SG [−A]-give PN   food
             'I gave Jama food'

        (b) *ni-'i-peɫe   Ja·ma kujá
             1SG.SUBJ-3SG [−A]-give PN   food

(The contrast in prefixes *m-* and *'i-* represents a difference in gender for the two nouns.)

(66)   immediate placement after the verb

        (a) ni-m-peɫe   Ja·ma kujá
             1SG.SUBJ-3SG [−A]-give PN   food
             'I gave Jama food'

        (b) *ni-m-peɫe   kujá Ja·ma
             1SG.SUBJ-3SG[−A]-give food PN

(67)   function as subject of the corresponding passive

        (a) Ja·ma ɸ-pela: kujá na: mi
             PN   3SG.SUBJ-give.PASS food by 1SG
             'Jama was given food by me'

        (b) *kujá i-pela   Ja·ma na: mi
             food 3SG.SUBJ-give.PASS PN   by 1SG
             'Food was given by me to Jama'

As we have defined it, argument structure is simply a listing of the number of a predicate's arguments with their precedence relations established by the

algorithm [+A] > [−A] > [ ]. Describing things in these terms entails that, for a ditransitive verb and for a canonical transitive verb (see chapter 3 by Andrews) the most prominent argument will necessarily be the [+A] participant, but for intransitive verbs it can either be a [+A] or [−A], depending on whether the verb is unergative or unaccusative, respectively. Just what grammatical constructions are sensitive to this notion of argument structure? When we look at the languages of the world we find that there are very many. In many languages, verb agreement will directly reflect the precedence relations of argument structure, so that the sole argument of an intransitive verb, regardless of whether it is [+A] or [−A], has the same morphological agreement form as the [+A] of a transitive verb, as in Iatmul of New Guinea (Staalsen (1972)):

(68)   (a) ntɨw yɨ-ntɨ   [+A] unergative
           man go-3SG.MASC
           'The man went'

       (b) ntɨw kɨya-ntɨ   [−A] unaccusative
           man die-3SG.MASC
           'The man died'

       (c) ntɨw takwə vɨ-ntɨ
           man woman see-3SG.MASC
           'The man saw the woman'

The verbal suffix *-ntɨ* agrees with the [+A] NP *ntɨw* 'men' in (68c); agreement with the [−A] NP *takwə* 'woman' is not possible. If the feminine agreement suffix for third person, *-lɨ*, were found on the verb, the sentence would have to mean 'the woman saw the man'; clearly the pronominal agreement is always with the [+A] participant of a transitive verb. The same agreement suffix *-ntɨ* is also found in (68a, b). The verb of (68a) is unergative, so again agreement is with a [+A] participant, but that of (68b) is unaccusative with a [−A] argument. Still the agreement is *-ntɨ* (3SG.MASC). The reason is that agreement in this language, unlike languages like Acehnese, is controlled by argument structure prominence. Because *kɨya-* 'die' is an intransitive verb, its most prominent argument (its sole argument) is a [−A], which thereby gets the same agreement forms as the highest ranked [+A] NP of unergative or transitive verbs. In unergative–unaccusative languages like Acehnese, on the other hand, verb agreement is not controlled by argument structure, but by the semantic parameters of perspective choice, [+A] or [−A], so that there are distinct sets of agreement clitics of [+A] and [−A] participants.

Nor is verb agreement the only grammatical construction controlled by argument structure prominence in Iatmul. Iatmul, like most Papuan languages (Foley (1986)), has a pervasive construction called clause chaining, in which

morphologically stripped-down verbs with their associated arguments follow
one after another in a sentence until a final fully inflected verb, as in this example:

(69)   ŋkəy-ət   yi-kə   waalə klə-laa   yə-kɨyə-ntɨ
        house-ALL go-DEP   dog   get-DEP   come-FUT-3SG.MASC
        'He will come after he has gone to the house and gotten the dog'

Crucially, such clause chaining patterns in Iatmul can only be formed when the
most prominent participant in the argument structure of each verb is shared, be
it [+A] or [−A]:

(70)   (a) vɨ-laa   yə-ntɨ
             see-DEP   come-3SG.MASC
             'He saw it and came'

        (b) vɨ-laa   kɨya-ntɨ
             see-DEP   die-3SG.MASC
             'He saw it and died'

The verb *vɨ-* 'see' is transitive in Iatmul, with two arguments, a [+A] and [−A]

(71)   *vɨ-* 'see'   someone sees someone/something

$$
\begin{array}{cc}
| & | \\
[+A] & [-A]
\end{array}
$$

vɨ- $<x, y>$

while *yə-* 'come' and *kɨya-* 'die' are unergative and unaccusative intransitive
verbs, respectively:

(72)

yə- 'come' someone moves to a place   *kɨya-* 'die' someone becomes dead

$$
\begin{array}{cc}
| & | \\
[+A] & [-A] \\
| & |
\end{array}
$$

yə- $<x>$ *kɨya-* $<x>$

Crucially, regardless of the [A] feature difference of the single arguments of
yə- 'come' and *kɨya-* 'die', clause chaining constructions are permissible with
them as long as coreference between them and the most prominent argument
of the other verb obtains:

(73)   (a) = (70a)

          [+A] [−A]   [+A]
           | | |
   vɨ - 'see' <x,   y>   *yə* - 'come' <x>
           | |
         3SG.MASC   3SG.MASC
         ↑_____↑
                 clause chaining OK

       (b) = (70b)

          [+A] [−A]   [−A]
           | | |
   vɨ - 'see' <x,   y>   *kɨya* - 'die' <x>
           | |
         3SG.MASC   3SG.MASC
         ↑_____↑
                 clause chaining OK
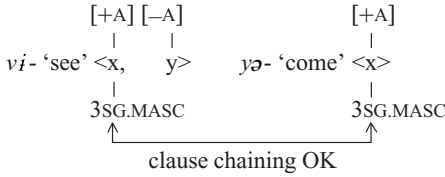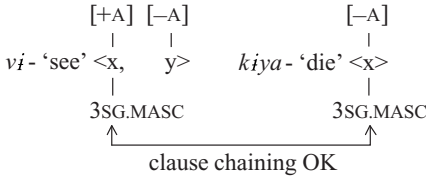
Patterns like Iatmal are not unique. Many languages have systems of clause linkage which are sensitive to coreference of the most prominent arguments in the linked verbs' argument structures. Control, in which the reference of a missing NP(ϕ) in a nonfinite clause is determined by an argument in the main finite clause, is in many languages typically a property of argument structure, namely the controlled NP is the most prominent NP in the argument structure of its governing verb. We see this in Yimas:

(74)   ┌──────────────────────┐
              ↓   ↓
       (a) [ϕ yampara-tu-wampuŋ] na-na-t-n
              stand.up-NFN-desire   3SG.S-PROG-feel-PRES
           'He feels like standing'

              ┌──────────────────────┐
              ↓   ↓
       (b) [ϕ mal-cu-wampuŋ] na-na-t-n
              die-NFN-desire   3SG.S-PROG-feel-PRES
           'He feels like dying'

              ┌──────────────────────┐
              ↓   ↓
       (c) [ϕ tpuk am-tu-wampuŋ] na-na-t-n
              sago eat-NFN-desire   3SG.S-PROG-feel-PRES
           'He feels like eating sago'
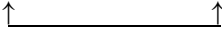
In (74) the nonfinite forms are complements of the verb *t-* 'feel' marked with the complementizer *wampuŋ-* 'heart, desire'. In all cases the highest argument of the argument structure of the complement is controlled by *na-* 3SG.S, the [−A] participant of 'feel'. The complement of (74c) contains a transitive verb *am-* 'eat' and its highest argument, the [+A] eater (ϕ), is missing and obligatorily

controlled, i.e. construed as equivalent to the [−A] of *t-* 'feel': the feeler is the one who wishes to eat. Both (74a, b) contain intransitive verbs, the unergative *yampara-* 'stand up' with a [+A] argument and the unaccusative *mal-* 'die' with a [−A] sole argument. In both cases, however, regardless of whether the argument is [+A] or [−A], it is controlled by *na-* 3SG.s, the [−A] experiencer of *t-* 'feel', because, as sole arguments, they are necessarily the highest argument in their respective argument structures. Control in Yimas, as with clause chaining in Iatmul, is directly sensitive to argument structure and the precedence relations therein.

2.2   *The notion of pivot*

Many languages – but, as we shall see, not all, nor perhaps most – have yet an additional level of syntactic organization beyond argument structure. This is the level of pivot organization. In a trivial sense, a pivot is any NP type which is the controller or the target of a particular grammatical construction. Thus, NPs which are realized through verb agreement in Iatmul, or controlled arguments in nonfinite complement clauses in Yimas, are pivots in this trivial sense. But this is not what we mean by a syntactic level of pivot organization. We follow Dixon (1979, 1994) in postulating the level of pivot organization in the grammar of a language, and thereby the syntactic notion pivot for it, only when the bulk of its grammatical constructions revolve around a particular NP type. English is a clear example of a language which exemplifies a pivot level of organization, for most of its syntactic processes centre around a single NP type, namely the subject, the NP that normally precedes the verb in a basic clause. Grammatical properties targeted by or targeting subjects include:
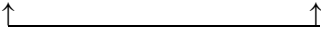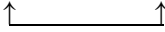
(75)   verb agreement
         (a) Ian was (SG) killing those chickens
         (b) *Ian were (PL) killing those chickens

(76)   preverbal position
         (a) Elizabeth loves Mary
         (b) Elizabeth loves too much
         (c) *Loves Elizabeth too much

(77)   controlled NP in nonfinite infinitive complements
         (a) Egbert wants [to φ take Mildred]

(b) *Egbert wants [Mildred to take ɸ]

(c) Egbert wants [to ɸ be taken by Mildred]
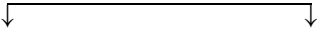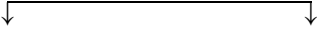
Example (77a) is fine because the subject of the complement is controlled by *Egbert*, but (77b) is ungrammatical because it is the non-subject [−A] participant which is controlled. In order to say something like (77b) grammatically, a passive construction (see section 4.1) must be used to realize the [−A] participant as subject, resulting in (77c). Similar considerations apply to other controlled nonfinite constructions like:

(78)   controlled NP in nonfinite participial relativization

(a) The guy [ɸ kissing the bartender] is my brother

(b) *The guy [the bartender kissing ɸ] is my brother

(c) The guy [ɸ being kissed by the bartender] is my brother

A final example of a construction targeting subjects concerns optional subject ellipsis of subjects of conjoined sentences with shared tense–aspect.

The sentences in (79) demonstrate that only subjects may be elided in these conjoined clauses.

(79)   (a) Fred walked to the library and ɸ met Bill

(b) *Fred walked to the library and Bill met ɸ

(c) Fred walked to the library and ɸ was met by Bill

(79b) is ungrammatical because a non-subject [−A] has been elided; again, a passive construction, as in (79c) in which the [−A] participant is realized as subject, restores grammaticality.

English subjects are clearly pivots in the sense in which we are using the term, and for cross-linguistic and theoretical purposes we will henceforth dispense with the term 'subject' and consistently use the term 'pivot' for this and related notions in other languages for the rest of this chapter. As with earlier perspective choices, this pivot notion will be captured through the use of features; a pivot participant will be marked [+pivot]; all other core arguments are [−pivot]. Thus, we have the following distribution of features:

(80)

```
                        ╱╲
                       ╱  ╲
                      ╱    ╲
              [−oblique]    [+oblique]
                 ╱╲
                ╱  ╲
               ╱    ╲
         [+pivot]   [−pivot]
```
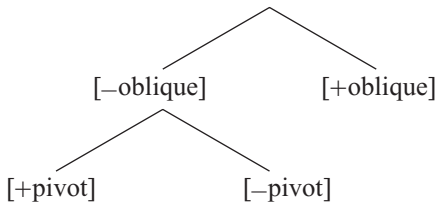
This set of features defines the array of grammatical relations for English and many other languages. Any participant which is assigned an [A] feature from its governing verb is also assigned the feature [−oblique]; otherwise it gets [+oblique]. This defines the first tier of grammatical relations and appears to be universal. Then, among the participants which have been assigned [−oblique], one is assigned [+pivot], the rest [−pivot]. The basis for this assignment in English and other languages is straightforward: assign [+pivot] to the most prominent argument in the argument structure. When there is more than one argument in the argument structure, as with transitive or ditransitive verbs, the most prominent argument is the one linked to the [+A] participant via the precedence algorithm [+A] > [−A] > [ ], so that the [+A] will be assigned [+pivot] and the sole or double [−A] argument(s) [−pivot]. Using this feature analysis, traditional English grammatical relations can be described as:

(81)   SUBJECT   OBJECT   OBLIQUE

$$\begin{bmatrix} -\text{oblique} \\ +\text{pivot} \end{bmatrix} \quad \begin{bmatrix} -\text{oblique} \\ -\text{pivot} \end{bmatrix} \quad [+\text{oblique}]$$
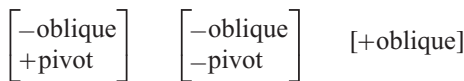
A sentence like

(82)   Egbert presented Mildred with an award

would have the following representation:

(83)
lexical semantic structure: *present* : someone causes someone to have something

```
                              |              |              |
                           [+A]          [−A]           [ ]
                              ╲            ╱              ╱
argument structure:           present <x, y, z>
                                      |
                            ╱         ╲         ╲
```
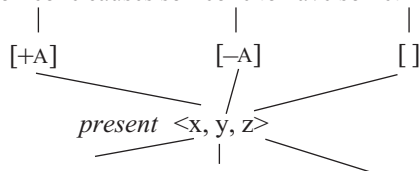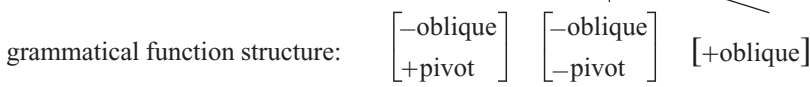
grammatical function structure: $\begin{bmatrix} -\text{oblique} \\ +\text{pivot} \end{bmatrix} \quad \begin{bmatrix} -\text{oblique} \\ -\text{pivot} \end{bmatrix} \quad [+\text{oblique}]$

All languages have grammatical functions because all languages have a contrast of core vs oblique NPs ([−oblique] vs [+oblique]). This is the minimal system of grammatical functions permissible and it does appear that there are languages in which this minimal system is all that there is. In other words, there are languages – indeed very many, perhaps the bulk of them – in which the concept of pivot as defined above, and therefore the features of [±pivot] for core arguments of their verbs, is entirely absent. Yimas is a very strong candidate for such a language. In example (74) above, we noted constructions like nonfinite complements in which the most prominent argument of an argument structure is the target for control, e.g. the [+A] argument of a transitive verb and the sole argument of an intransitive verb, be it unergative ([+A] argument) or unaccusative ([−A] argument). There are also constructions in Yimas, however, in which the [−A] argument of a transitive verb or the sole argument of an intransitive or unergative unaccusative verb is the target of the construction – for example, the scope of elevational/directional affixes:

(84) (a) kay   i-na-l-ampu-n
          canoe.VIII.SG VIII.SG.S-PROG-down-float-PRES
                              [−A]
          'The canoe is floating down there'

      (b) ka-mpu-tra-ya-n
          LIKELY-3PL.S-about-come-PRES
                  [−A]
          'They can come about'

Elevational/directional affixes indicate the spatial coordinates of the conceptual event described by the clause, notions like 'up', 'down', 'in', 'out', 'about', etc. With intransitive verbs it is the location/direction of the single argument which is predicated, but with transitive verbs it is always that of the [−A] participant, as the transitive equivalents of (84) demonstrate:

(85) (a) kay   nan-l-arm-na-ŋkan-i
          canoe.VIII.SG IMPER.PL-down-board-IMPER-PC-VIII.SG
                          [+A]   [−A]
          'You few board the canoe down below'
          *'You few down below board the canoe'

      (b) nmpi   ay-cra-wampak-ɲa-ŋkt
          leaf.VII.PL HORT.PL-about-throw-IMPER-PC
          [−A]   [+A]
          'Let us few send messages about'
          *'Let us few about send messages'

In both cases it is the location of the [−A] participant which is specified by the elevational/directional affix; the canoe is down in (85a) and the messages are to be sent about in (85b). The single argument of (84) corresponds to the [−A] argument of the transitive verbs in (85) and both are under the scope of the elevational/directional prefixes. There seems to be no coherent notion of pivot in these Yimas data; consider for illustration, just transitive verbs: control in nonfinite complements targets [+A] arguments, while scope of elevational/directional affixes targets [−A] arguments.

But in fact the picture is even more complicated. Control in nonfinite complements and elevational/directional scope in Yimas are unusual in showing a preference among core arguments, albeit different core arguments. The large majority of grammatical constructions in Yimas show no such preferences: any core argument, regardless of its function, is a potential target. A good example is clause chaining. Unlike Iatmul and most other Papuan languages (see (73)), Yimas does not restrict clause chaining to situations of coreference among the most prominent arguments in the argument structure. Dependent verb clause chains can be formed as long as there is a coreferential core argument:

(86)   (a) tmal   kray-mpi ya-kay-am-wat   amtra
             sun.V.SG dry-DEP   V.PL.P-1PL.A-eat-HABIT food.V.PL
                                 [−A]   [+A]
             'The sun having dried it, we always eat the food'

        (b) panpan-tat-mpi   mnta narmaŋ
             pound sago-start-DEP then woman.II.SG

             ŋka-pu-k-mp-n   pia-n-i-k-nakn
             go by land-toward-IRR-VII.SG-OBLIQ talk P-3SG.A-tell-IRR-3SG.DU
                                 [−A] [+A]   [−A]
             '(He) starting pounding sago and then (his) wife comes toward
             him, she tells him'

In (86a) the argument shared by the two clauses is *amtra* 'food', functioning as [−A] in both clauses. For (86b) the shared argument between the dependent clause and the final independent clause is the understood *panmal* 'male, husband'. This functions as the sole [+A] argument in the first dependent clause with the intransitive unergative verb *panpan-tal-* 'start pounding sago', but as a [−A] participant in the final independent clause. Clearly, these clause chaining constructions are much more freely formed in Yimas than in Iatmul – indeed, on any shared core argument.

Investigation of many other Yimas constructions would force similar conclusions. The fundamental principle organizing Yimas grammar is the contrast between the grammatical functions [−oblique] versus [+oblique]. Oblique

participants are necessarily case-marked with the oblique case suffix *-n* ∼ *-nan* and never permitted to exhibit verbal pronominal agreement. Core [−oblique] grammatical functions, whether [+A] or [−A], may never be case marked at all and typically exhibit verb agreement. Beyond that, core arguments show little in the way of systematic patterns of contrast in syntactic behaviour so as to establish a pivot notion in the language. We seem to be on safe ground in concluding that the basic organization of grammatical functions in Yimas revolves around the features [±oblique], and that [±pivot] is not employed in the language. We may also like to propose that Yimas is hardly unique; a careful study of the syntactic structure of languages will probably demonstrate numerous examples of the Yimas pattern.

2.3   A typology of pivots

Languages can differ as to their basic choice of pivot. In English, as we saw above, the basic choice of pivot for a canonical transitive or ditransitive verb is [+A] (see chapter 3 by Andrews). With intransitive verbs, of course, the choice is necessarily the single argument, so the choice for English could be summarized as:

(87) $\boxed{\begin{array}{l} [+\text{A}] \\ \\ [\pm\text{A}] \end{array}}$ $\begin{array}{l} V_{\text{TRANS}} \quad [-\text{A}] \\ \\ V_{\text{INTR}} \end{array}$

[+pivot]   [−pivot]

Note that the [+pivot] [−pivot] contrast here is basically the same as the English pronominal case-marking system:

(88) $\boxed{\begin{array}{l} \text{I} \\ [+\text{A}] \\ \\ \text{I} \\ [+\text{A}] \\ \\ \text{I} \\ [-\text{A}] \end{array}}$ $\begin{array}{l} \text{followed} \quad \text{him} \\ \qquad\qquad [-\text{A}] \\ \\ \text{ran} \\ \\ \\ \text{fell} \end{array}$

Nominative   Accusative

English has a single case form, the nominative, which marks the [+A] argument of transitive verbs, the [+A] argument of intransitive unergative verbs and the [−A] argument of intransitive unaccusative verbs, and contrasts this with an accusative case which marks the [−A] argument of transitive verbs. Nominative case is the case of the [+pivot] [−oblique] argument, and accusative the case of the [−pivot] [−oblique] argument (also the oblique NPs, as complements of their governing prepositions: *we left with him*). We may succinctly describe the pivot structure of English as nominative–accusative. As we have seen, English does possess constructions, namely passives, which do present [−A] arguments of basic transitive verbs as the pivot and hence in the nominative, but crucially such structures are intransitive and therefore are essentially just intransitive unaccusative verbs. Their intransitivity is clearly demonstrated by the fact that no [+A] argument is necessary in passive constructions:

(89)

I
[−A]
[+pivot]
NOM

was followed/kissed/hit/seen

Other languages with nominative pivot choices ([+A] for transitive verbs and [±A] for intransitive verbs) include German (Foley and Van Valin 1984), French, Italian, and possibly many Bantu languages like Chichêwa and Swahili.

There are, however, two other systems of pivot choices. One of these is the mirror image of the English system. As illustrated in (88), English pronominal case marking is nominative–accusative, contrasting a case for the [+A] argument of a transitive verb and the single [±A] argument of an intransitive verb (nominative) with that of the [−A] argument of a transitive verb (accusative). Many languages have a case-marking system which contrasts the [−A] argument of a transitive verb and the [±A] argument of an intransitive verb against the [+A] argument of a transitive verb. Such systems are called ergative–absolutive case-marking systems and can be illustrated with these examples from the Australian language Dyirbal (Dixon (1972)):

(90)

(a)

balan	ḍugumbil	baŋi-ɲu
DET.ABS	woman.ABS	come-TNS
	[+A]	

'The woman came'

(b)

balan	ḍugumbil	baḍi-ɲu
DET.ABS	woman.ABS	fall-TNS
	[−A]	

'The woman fell'

(c)

balan	ḍugumbil	baŋgul	yaṛa-ŋgu	balga-n
DET.ABS	woman.ABS	DET.ERG	man-ERG	hit-TNS
	[−A]		[+A]	

'The man hit the woman'

ABSolutive   ERGative

Unlike many Australian languages which are pivotless, Dyirbal does have a pivot, but one that is the opposite choice to English and parallels its ergative–absolutive case-marking system: the basic pivot of a transitive verb in Dyirbal is the [−A] argument or the absolutive NP, so Dyirbal absolutive case is [+pivot], [−oblique], while ergative is [−pivot], [−oblique]. Parallel to English passive, Dyirbal also possesses a construction, called the *antipassive*, which allows the [+A] argument of a transitive verb to assume the [+pivot] feature, but again this is an intransitive construction, this time unergative, not requiring a [−A] participant, which, if present must be in a [+oblique] case like dative. The antipassive version of (90c) is (91) (-*ŋa*(*y*) is the derivational verbal suffix for antipassive):

(91)   bayi   yaṛa   (bagun   ḍugumbil-gu)   balgal-ŋa-ɲu
       DET.ABS   man.ABS   DET.DAT   woman-DAT   hit-ANTIPASS-TNS
           | |
         [+A]   [−A]
           | |
       $\begin{bmatrix} -\text{oblique} \\ +\text{pivot} \end{bmatrix}$ [+oblique]

       'The man hit the woman'

   The pivot system of Dyirbal is ergative–absolutive (ergative: [−oblique], [−pivot]; absolutive: [−oblique], [+pivot]), as opposed to English, which

is nominative–accusative (nominative: [−oblique], [+pivot]; accusative: [−oblique], [−pivot]). This claim can be substantiated for Dyirbal in the same way it was for English: by showing a number of grammatical constructions for which the [−A] pivot is the controller or target. A number of these constructions have English analogues; consider, for example:

(92)   controlled NP in nonfinite purposive clauses

(a) balan   ḍugumbil   baŋgul   yaṛa-ŋgu balga-n [ɸ baḍi-gu]
             [−A]   [+A]   [−A]
    DET.ABS woman.ABS DET.ERG man-ERG hit-TNS fall-PURP
    'The man hit the woman, (causing her) to fall'

(b) *balan   ḍugumbil   baŋgul   yaṛa-ŋgu wawu-n [ɸ
             [−A]   [+A]   [+A]
    DET.ABS woman.ABS DET.ERG man-ERG fetch-TNS
    balan   nayinba   walmbil-i
             [−A]
    DET.ABS girl.ABS   get up-PURP
    'The man fetched the woman to get the girls up'

(c) balan   ḍugumbil   baŋgul   yaṛa-ŋgu wawu-n [ɸ
             [−A]   [+A]   [+A]
    DET.ABS woman.ABS DET.ERG man-ERG fetch-TNS
    bagun   nayinba-gu walmbil-ŋay-gu
             [−A]
    DET.DAT girl-DAT   get up-ANTIPASS-PURP
    'The man fetched the woman to get the girls up'

                                        Dixon (1972, 1994)

Remember the pivot for Dyirbal is the absolutive NP, the [−A] argument of a transitive verb or the sole [±A] argument of an intransitive verb. The pivot is the target for control in these nonfinite purposive clauses. Example (92a) is fine because the controlled NP is the [−A] sole argument of an unaccusative intransitive verb *baḍi-* 'fall', which, if overt, would be in the absolutive case (see (90b)). Example (92b), on the other hand, is ungrammatical because the targeted NP for control is the [+A] argument of a transitive verb *walmbil-* 'get up', which would appear, if overt, in the [−pivot] ergative case (note that the English translation of (92b) is perfectly grammatical because English, of course, has the opposite choice for pivot from Dyirbal, namely the [+A] argument of a transitive verb). For the meaning of (92b) to be expressed grammatically, an

antipassive must be used, in which the [+A] argument is presented as the pivot of a derived intransitive unergative verb. This is found in (92c): the transitive verb root *walmbil-* 'get up' is now marked with the detransitivizing antipassive suffix *-ŋa(y)*. Its core [+A] argument, if overt, would have to be marked with absolutive [+pivot] (see (91)). As pivot, it can now be a target for control so that (92c) is well formed.

Languages with an ergative–absolutive alignment for pivot choices, i.e. the pivot is the [−A] participant of a transitive verb in contrast to the [+A] choice for languages with a nominative–accusative alignment, were once thought to be rare. But intensive study since the mid-1970s has unearthed more examples. Besides Dyirbal and some other languages of far north Queensland in Australia (see Dixon (1977a, 1980, 1981, 1994)), other languages of this type include Eskimo (Manning (1996)), some Mayan languages, especially those of the Mamean subgroup like Mam (England (1983a, 1983b, 1988)) and perhaps a number of Western Austronesian languages like Chamorro (Cooreman (1987, 1988)). An interesting problem confronted by all languages with this typology is the class of ditransitive verbs, with their multiple [−A] participants. There can only be one [+pivot] NP per clause. Ditransitive verbs present no problems for nominative–accusatively aligned [+pivot] languages like English because the pivot is the [+A] participant, of which there is only one:

(93)   Egbert gave   Hortense a   python
          | | |
        [+A]   [−A]   [−A]
          | | |
$\begin{bmatrix} -\text{oblique} \\ +\text{pivot} \end{bmatrix}$ $\begin{bmatrix} -\text{oblique} \\ -\text{pivot} \end{bmatrix}$ $\begin{bmatrix} -\text{oblique} \\ -\text{pivot} \end{bmatrix}$

But true ditransitive verbs should be impossible or at least highly shunned in ergative–absolutively aligned [+pivot] languages, because it would potentially permit two pivots which is impossible, so that structures like the following should be ungrammatical:

(94)   *Egbert gave   Hortense a   python
           | | |
         [+A]   [−A]   [−A]
           | | |
$\begin{bmatrix} -\text{oblique} \\ -\text{pivot} \end{bmatrix}$ $\begin{bmatrix} -\text{oblique} \\ +\text{pivot} \end{bmatrix}$ $\begin{bmatrix} -\text{oblique} \\ +\text{pivot} \end{bmatrix}$
                               ↑_____*_____↑

In language after language of this typology, this prediction is borne out: there are no true ditransitive verbs; there is only one [−A] argument so that 'give' is a transitive verb only:

(95)   Dyirbal

    (a) balam   miraɲ   baŋgul   yaṛa-ŋgu
                  [−A]   [+A]
           DET.ABS   bean.ABS   DET.ERG   man-ERG
           wuga-n   bagun   ḍugumbil-gu
                                [+oblique]
           give-TNS   DET.DAT   woman-DAT
           'The man gave beans to the woman'

    (b) balan   ḍugumbil   baŋgul   yaṛa-ŋgu
                    [−A]   [+A]
           DET.ABS   woman.ABS   DET.ERG   man-ERG
           wuga-n   baŋgum   miṛaɲ-ḍu
                                [+oblique]
           give-TNS   DET.INSTR   bean-INSTR
           'The man gave the woman (with) beans'   Dixon (1972)

Dyirbal *wuga-* 'give' is like English *spray*, a simple transitive verb with variable [−A] assignment, either to the gift, with the recipient in [+oblique] dative case, or to the recipient, with the gift in the [+oblique] instrumental case. Other languages of this typology show a similar restriction of 'give' to this transitive class only:

(96)   (a) Mam

        ma-a^ʔ   ɸ-tsaj   ky-q'o-^ʔn   pwaq   q-ee

        TNS-EMPH   3SG.ABS-DIR   3PL.ERG-give-DIR   money   1PL-DAT
                        | | |
                  [−A]   [+A]   [ ]
                    | | |

           $\begin{bmatrix} -\text{oblique} \\ +\text{pivot} \end{bmatrix}$ $\begin{bmatrix} -\text{oblique} \\ -\text{pivot} \end{bmatrix}$ [+oblique]

        'They gave the money to us'   England (1983a)

(b) Chamorro

ha-na'i	hao	si	Juan	ni	lep blo
3SG.ERG-give	2SG.ABS	DET	John	OBLIQ	book

ha-na'i   hao   lep blo

[+A]   [−A]   [ ]

$$\begin{bmatrix} -\text{oblique} \\ -\text{pivot} \end{bmatrix} \quad \begin{bmatrix} -\text{oblique} \\ +\text{pivot} \end{bmatrix} \qquad [+\text{oblique}]$$

'John gave you (with the book)'   Cooreman (1988)

In both Mam and Chamorro only one [−A] NP is found with 'give', but which one it is differs. In Mam the gift is [−A] and hence [+pivot], while in Chamorro it is the recipient. Thus, in Mam it is the recipient which is [+oblique], but in Chamorro, the gift.

In addition to pivot choices we have seen thus far – [+A] argument of transitive verbs in nominative–accusative languages like English, [−A] of transitive verbs in ergative–absolutive languages like Dyirbal, neither as in pivotless languages like Yima – there is a fourth logical possibility: both [+A] and [−A] of a transitive verb are available to [+pivot] status, but, of course, due to the one pivot per clause constraint, only one choice can be made in any one clause. Such languages could be termed *symmetrical pivot languages*, in contrast to *asymmetrical pivot languages* like English and Dyirbal. In symmetrical languages there is no strong preference for [+A] or [−A] to be the [+pivot] NP of the clause; either can be, according to wider syntactic or textual constraints. Philippine languages, of which Tagalog is a good example, illustrate symmetrical languages. In Tagalog [−oblique] arguments are marked with the preposition *ng* and [+oblique] arguments with *sa*. The pivot NP has the preposition *ang* and the verb takes a series of affixes to indicate the [±A]-feature status of the pivot (-*um*- 'pivot = [+A]' is an infix which occurs between the initial consonant of the root and the following vowel):

(97) (a) b-um-ili   ang   lalake ng   isda

pivot=[+A]-buy [+pivot] man [−oblique] fish

[+A]   [−A]

$$\begin{bmatrix} -\text{oblique} \\ +\text{pivot} \end{bmatrix} \qquad \begin{bmatrix} -\text{oblique} \\ -\text{pivot} \end{bmatrix}$$

'The man bought fish'

(b) bi-bilh-in   ng   lalake ang   isda

FUT-buy-pivot=[−A] [−oblique] man   [+pivot] fish

$$
\begin{array}{cc}
[+\text{A}] & [-\text{A}] \\
\Big| & \Big| \\
\begin{bmatrix} -\text{oblique} \\ -\text{pivot} \end{bmatrix} & \begin{bmatrix} -\text{oblique} \\ +\text{pivot} \end{bmatrix}
\end{array}
$$
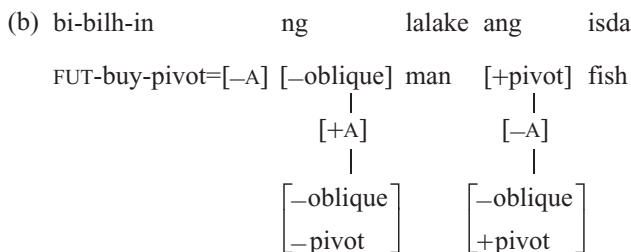
'The man will buy the fish'

Note that either the [+A] or [−A] argument can be pivot, the choice being signalled by different affixation to the verb, the infix *-um-* for '[+A] = pivot' and the suffix *-in* for, '[−A] = pivot'. Note further that, regardless of which argument becomes pivot, the other argument remains [−oblique] and marked by *ng*. This is in stark contrast to asymmetrical languages like English or Dyirbal in which marked constructions like passive or antipassive must be used when anything other than the normal pivot choice is made, forcing detransitivization of the clause, with the erstwhile pivot appearing as a [+oblique] NP or disappearing from the clause entirely. Clearly, languages like English or Dyirbal are asymmetrical: one argument type is strongly favoured to be pivot, and if it fails to assume this function, it ceases to be [−oblique]. This is not the case in symmetrical languages like Tagalog: either choice is possible, with no major disruption to clause structure regardless of which is chosen.

Tagalog pivots have many of the same properties we have come to expect of them from English or Dyirbal, for example, controlled NP in relativization (remember *-um-* 'pivot = [+A]' is an infix):

(98)   (a) lalake-ng [b-um-ili   φ   ng   isda]
             man-REL pivot=[+A]-buy [+A] [−oblique] fish
             'The man who bought fish'

    (b) *isda-ng [b-um-ili   ang   lalake φ]
             fish-REL pivot=[+A]-buy [+pivot] man   [+A]
             'The fish the man bought'

    (c) isda-ng [bi-bilh-in   ng   lalake φ]
             fish-REL FUT-buy-pivot=[−A] [−oblique] man   [−A]
             'The fish the man will buy'
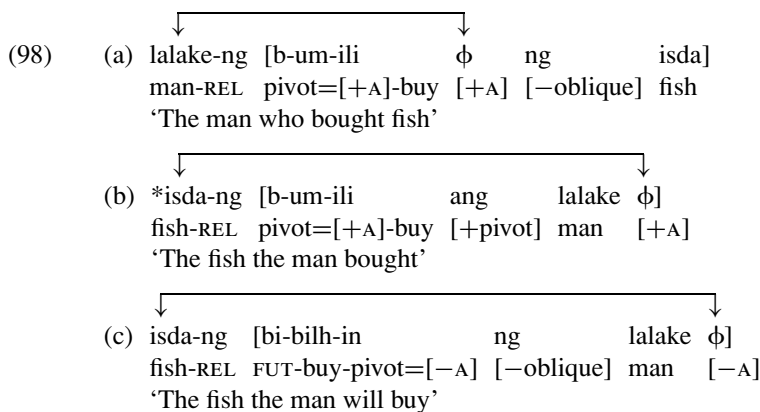
Example (98a) is grammatical because the controlled NP in the relative clause is the pivot, identified as being the [+A] argument by the infix *-um-*. Example (98b) is ungrammatical because the controlled NP and the pivot are not identical: the pivot is the [+A], again identified by the infix *-um-*, but the controlled NP is the [−A] argument *isda* 'fish'. In order to form such a relative clause, the [−A] must be made pivot, as it is in (98c); the verb is suffixed with *-in* indicating the [−A] argument is the pivot and therefore is a proper target for control.

The following table summarizes the typology of pivots developed in this and the last section.
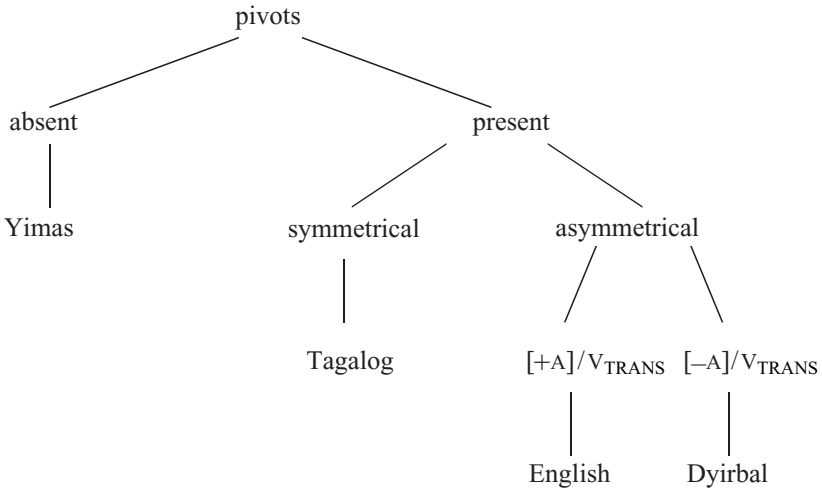
3   On information structure

3.1   *The discourse status of noun phrases*

Argument structure, the precedence relations among constituents, the presence or absence of pivot structure, whether pivot choice is symmetrical or asymmetrical, etc. are all aspects of the basic lexically determined syntactic organization of a clause. Most sentences do not occur in isolation but are uttered in the course of an ongoing verbal interaction, e.g. as part of an ongoing story, in response to an earlier question or assertion, as the next step in a Thai curry recipe, etc. Thus, speakers can assume that there is information about already mentioned events and participants common to both them and their addressees, but other information which is not. The structuring of sentences along these

parameters – that is, the discourse-given information-carrying status of sentence constituents – is called *information structure*. The operation of information structure is well illustrated by mini-dialogues, like these drawn from Russian:

(99)   (a) Q: któ   zaščiščajet Víktor-a
                      who.NOM defends   Victor-ACC
                      'Who defends Victor?'
            A: Víktor-a   zaščiščajet Maksím
                      Victor-ACC defends   Maxim.NOM
                      'Maxim defends Victor'

        (b) Q: Kogó   zaščiščajet Maksím
                      who.ACC defends   Maxim.NOM
                      'Whom does Maxim defend?'
            A: Maksím   zaščiščajet Víktor-a
                      Maxim.NOM defends   Victor-ACC
                      'Maxim defends Victor'   Comrie (1987: 95)

Note that the answer to the questions in English can have exactly the same word order; this is not possible in Russian. Rather the NP which provides the answer to the question word *któ* 'who' or *kogó* 'whom' must always occur clause-finally, while the information already established by the question must precede it. The question thus sets up certain information expectations of NPs that must be realized in the information structure of the answer. The NP providing the answer to the question word is the *focus* of the clause, expressing the new information the clause is expected to provide. The whole point of uttering it in first place is to register this new information in the addressee's store of knowledge. The remainder is what is assumed by the speaker to be common knowledge shared by himself and his addressee; it is *presupposed*. The information structure for (99a) could be represented as:

(100)   Q: Presupposed: someone is defending Victor
             Focus: who is that someone?

        A: Presupposed: someone is defending Victor
             Focus: that someone is Maxim

With the exception of question words like *któ* 'who' and *kogó* 'whom', which always come first even though they represent focussed information, Russian has a fairly rigid rule of information structure: presupposed information precedes focussed information. Focussed information typically occurs at the end of a sentence:

(101)   Presupposed: someone is defending Victor

Focus: that someone is Maxim

Víktor -a	zaščiščajet	Maksím
Victor-ACC	defends	Maxim.NOM
PRESUPPOSED		FOCUS

Example (101) presents an obvious question as to the information status of the first NP, *Víktor-a* 'Victor-ACC'. This corresponds to the *topic* of the sentence, another elemental notion of information structure, again best illustrated through the use of mini-dialogues:

(102)   S: Maksím   ubivájet Alekséj-a
             Maxim.NOM kills   Alex-ACC
             'Maxim kills Alexei'

Q: a   Víktor-a
   and Victor-ACC
   'and Victor?'

A: Víktor-a   Máksim   zaščiščajet
   Victor-ACC Maxim.NOM defends
   'Maxim defends Victor'   Comrie (1987: 96)

Typically, a sentence expresses a comment about some entity. This entity is really what the sentence is about, its presupposed starting point. It is referred to by an NP which corresponds to the *topic* of the sentence. The topic is the link which ties the information communicated in this sentence with what has preceded. It is the source of coherence which makes the sentence relevant and interpretable within the context of the ongoing verbal interaction. The topic is fundamentally the presupposed information of the sentence that the remainder comments upon. In Russian, as in many languages, the topic typically occurs clause-initially; the initial NP in the final answer in the above mini-dialogue has the following topic–comment structure.

(103)   Topic: concerning Victor
             Comment: Maxim defends him

There are, then, two systems expressed in the information structure of the Russian clause, that of presupposed–focus and topic–comment. The information structure of the final answer in (102) could be diagrammed as:

(104)   topic   ‖   comment
             Viktor-a ‖ Máksim ‖ zašc̆isc̆ajet
                  presupposed ‖   focus

The basic word order rules that realize information structure in Russian stipulate that topic precedes comment and presupposed precedes focus. This typically results in a clausal order with the topic NP sentence-initially and the focus NP clause-finally. The topic prototypically denotes a presupposed established entity, while the focus supplies part of the comment, newly provided information, on that entity. A Russian sentence then prototypically proceeds from what is known, already established and presupposed in the topic, to what is unknown and only now supplied as new information by the focus.

Information in English is realized in much the same way as Russian, through word order, in spite of the fact that English word order, unlike Russian, is also determined by lexical constraints, i.e. a [+A] argument precedes a transitive verb while a [−A] follows: *Egbert killed Ian* has assignments of [+A] and [−A] that are exactly opposite to *Ian killed Egbert*. Still, the basic principle for English is the same as for Russian: topic occurs sentence-initially, and focus, finally. This principle shows up clearly because not all English word order is rigidly fixed; in particular [+oblique] NPs have greater freedom; e.g. *the children are playing in the yard* versus *in the yard the children are playing*. These realize different information structures, as is apparent from constructing mini-dialogues:

(105)   (a) Q: Where are the children playing?
              A: The children are playing ‖ in the yard
                   PRESUPPOSED ‖ FOCUS
              A: ?In the yard ‖ the children are playing
                   FOCUS ‖   PRESUPPOSED

         (b) Q: Are the children singing in the yard?
              A: No, in the yard the children ‖ are playing
                   PRESUPPOSED ‖ FOCUS

So we find that the normal position for focussed information in English is sentence-finally, as in Russian. However, as we have seen, English does have lexically specified argument structure constraints on its word order options for [−oblique] NPs, so that if the focussed NP is [−oblique] there may be deviations from this rule. The language does, however, require that such deviations be signalled, and this is normally done by a marked high falling pitch on a focussed NP which is not in normal sentence-final position:

(106) Q: Who saw Bill?

       A: Bill was seen   ‖   by John
            PRESUPPOSED   ‖   FOCUS

       A: John   ⌒   ‖   saw Bill
              FOCUS   ‖   PRESUPPOSED

There are two opposing word order principles in these answers: one, from argument structure, that specifies that the [+A] argument as [+pivot] proceeds a transitive verb and a [−A] argument follows; and another, from information structure, that stipulates the order 'presupposed precedes focus'. The problem is the focussed [+A] NP. In the first answer, the information structure principle wins out; the focussed NP occurs clause-finally, but, because it is a [+A] participant, a passive construction which realizes it as [+oblique] (see section 4) is necessary to do this. In the second answer the lexical argument structure principle wins out, so that the [+A] argument is realized as [−oblique] [+pivot]. This violates the information structure principle because the focus NP now occurs clause-initially. In order to indicate this marked position for focus, a high falling pitch occurs on the focus NP.

   The preverbal, prototypically clause-initial position of the [+pivot] NP in English also complicates the realization of the topic notion in this language. There is in fact a strong correlation between the concepts of pivot and topic in English and other languages, a relationship I will discuss in greater depth in section 3.3. A typical way to express alternatives of topic choice in English is to select different pivots, for example:

(107) (a) Tears   ‖   streamed down her face
             TOPIC   ‖   COMMENT

       (b) Her face   ‖   streamed with tears
             TOPIC   ‖   COMMENT

(108) (a) Blood   ‖   flowed in the streets
             TOPIC   ‖   COMMENT

       (b) The streets ‖   flowed with blood
             TOPIC   ‖   COMMENT

The verbs *stream* and *flow* are ambiguously unergative or unaccusative intransitive verbs; as such they can select either [+A] or [−A] sole arguments with

a corresponding semantic difference: moving object [+A] versus fully affected location [−A]. Correlated to this choice, however, is topic selection; whichever [+A] or [−A] is [+pivot], it is also topic. In other similar cases, this option is tied to the selection of a different verb which lexicalizes pivot and hence topic choice:

(109) (a) Q: Where is   the dot?
              TOPIC   ‖   COMMENT

              A: The dot ‖ is ‖   inside the circle
              PRESUPPOSED   ‖   FOCUS

              A: ?The circle surrounds the dot

       (b) Q: Where is the circle?

                   TOPIC   ‖   COMMENT
              A: The circle ‖   surrounds ‖   the dot
                                                 FOCUS
                   PRESUPPOSED
              A: ?The dot is inside the circle

The question in (109a) sets up the dot as presupposed and topic and then questions its location, which thereby must be the focus of the answer. The first answer with *dot* as topic and pivot and *inside the circle* in clause-final focus position demonstrates proper English information structure in answer to this question. The second answer in which the transitive verb *surround* selects *the circle* as pivot is decidedly odd, because pivots are normally linked to topic NPs, but *the circle* is in fact focus. Now, consider question (109b) which presents *the circle* as presupposed and topic and questions its location, setting this up as focus of the following answer. Now, the choice of the transitive verb *surround* is perfectly acceptable, because *the circle* is pivot and unmarked topic, while *the dot*, the landmark which specifies the location of the circle, follows the verb and can function as focus in normal clause-final position. Thus, the information structure of the first answer is expected, while that of the second is again odd because it presents the focus as pivot and unmarked topic and the topic in the clause-final position of the typical focus NP. We can summarize the typical unmarked linking of English lexical-argument structure and information structure as:

(110)

$$
\begin{array}{ccc}
\text{NP} & \text{V}_{\text{TRANS}} & \text{NP} \\
| & & | \\
[+\text{A}] & & [-\text{A}] \\
| & & | \\
\begin{bmatrix} -\text{oblique} \\ +\text{pivot} \end{bmatrix} & & \begin{bmatrix} -\text{oblique} \\ -\text{pivot} \end{bmatrix} \\
| & & | \\
\text{TOPIC} & & \text{FOCUS}
\end{array}
$$
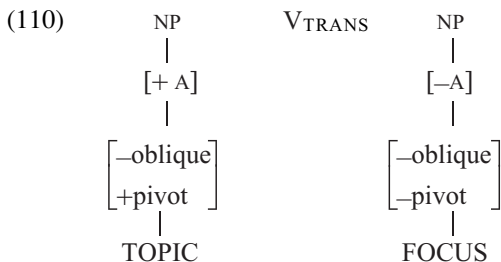
The linkings of (110) are, of course, not fixed; English possesses many devices which allow alternative construals of lexical argument structure and information structure, for example intonation and high falling pitch in NPs, as in (106). Any syntactic construction which affects pivot choice can also affect topic choice. The most pervasive construction in English which affects pivot choice is the passive (discussed in detail in section 4.1) and the contrast between active and passive is typically tied to differences in information structure:

(111)   (a) The manager || sacked the workers
                 TOPIC || COMMENT
        (b) The workers || were sacked by the manager
                 TOPIC || COMMENT

Again, this information structure contrast can be most easily viewed in constructed mini-dialogues:

(112)   (a) Q: Who saw Bill?

             A: Bill || was seen by John
                TOPIC || COMMENT
        (b) Q: Whom did Bill see?

             A: ?John || was seen by Bill
                TOPIC || COMMENT

The first pair exhibits normal information structure in the answer: the [−A] participant, *Bill*, is presupposed and topic as established by the question, so that an answer in the passive voice with the [−A] as pivot and topic is fine. Example (112b), however, is decidedly bizarre, because the question established *Bill*, the [+A] participant, as the topic. Thus, an answer in the passive as here, which presents the [−A], *John*, as topic and the [+A] *Bill* as focus, is a

completely unexpected information structure. Rather, an active sentence should be used: *Bill*, the [+A], as pivot and topic, and *John*, the [−A], as focus: *Bill saw John*.

While the notions of pivot and topic are closely correlated in English, they are not isomorphic: there are clear cases of topics which are not pivots. One of the most common constructions of this type presents topics occurring sentence-initially and preceding pivots (these will be discussed in greater detail in sections 3.3 and 5):

(113)   TOPIC   PIVOT

   (a) Last night   │││   I   │││   saw three movies in town
   (b) As for John,   │││ Ian   │││   is in love with him
   (c) Soukous,   │││   I   │││   think it's the greatest African
                                        twentieth-century contribution to
                                        civilization

These constructions typically express topics which are not highly salient right at the moment when they are mentioned in the discourse, but are nonetheless being highlighted; they are, however, presupposed:

(114)   Q: What's your favourite dance music?

         A: Soukous ││ my feet find irresistible
            TOPIC   ││   COMMENT

Soukous, as a type of African dance music, while not explicitly mentioned in prior discourse, is presupposed by the general cover term *dance music* in the question.

3.2   *The information status of noun phrases*

As we have seen, topics are typically presupposed information, the starting point of the sentence. Focussed NPs, on the other hand, are the end goal of the sentence, the information which the speaker intends to introduce into the discourse. Topics are closely correlated with given or old information, which is currently in the frame of the discourse, while focussed constituents are new information, just being introduced into the discourse. The concept of given information is more or less equivalent to presupposed, but new information need not (though it usually does) correspond to the focussed constituent:

(115)   Q: What happened?

          A: An earthquake just hit Los Angeles

Arguably, all the information in the answer is new, but only the final NP, *Los Angeles*, is really the focus, as demonstrated by its high falling pitch. This analysis differs from that in Lambrecht (1994), who argues that answers like (115) represent what he calls 'sentence focus' structure, in which the whole sentence is focussed information. That is a possible alternative, but I prefer to distinguish between non-presupposed, asserted information and focussed information, which is the particular new, asserted information the speaker wishes to highlight. In the answer above, that is just *Los Angeles*, with its high falling pitch; the remainder of the sentence is new, but unfocussed, information.

Other types of mismatches between topic and presupposed/given and between focus and new also occur. A particularly common type is contrastive NPs. These are marked as focus by high falling pitch, but are typically presupposed:

(116)   Considering the suspects, only Egbert has a motive

Egbert is one of several suspects in a criminal investigation; he is presupposed, yet new information is provided about him, as focus, as being the only person among them who has a motive. He is set up in contrast to the other suspects, presupposed, but focussed. There are also examples of topics which are new information. Perhaps the best-known examples of these are found in the opening lines of stories, such as fairy tales:

(117)   Once upon a time, a little old man lived in a big house on the hill

A little old man is the topic of this sentence, but, as it initiates a discourse, it is obviously not presupposed, given information, but new.

Besides given versus new, NPs may bear other kinds of information statuses, depending on the speaker's view of her addressee's current knowledge. An NP is referential if the speaker intends that it refer to a particular entity in the world. For example, if someone rings me up and asks me *what are you doing?* and I respond *I'm looking for a snake*, the NP *a snake* could be either referential or nonreferential. If I have a particular snake in mind, say, my pet diamond python, then *a snake* would be referential, as in

(118)   I'm looking for a snake. Carbon escaped from his cage

where Carbon is the pet diamond python's name. If, on the other hand, I'm lonely and just want a pet snake, then *a snake* would be nonreferential, as in:

(119)   I'm looking for a snake for a pet; what kinds do you sell?

Pronouns are typically referential, but at least two in English, *you* and *it*, have common nonreferential uses:

(120)   (a) You pay your money and you take your chances
         (b) It seems to me that you've rather overspent your account

The information status of NPs can further be classified as 'definite' or 'indefinite'. An NP is definite when the speaker presupposes the addressee can uniquely identify its referent from the universe of discourse; otherwise it is indefinite. So if I come up to you and announce *Egbert bought the snake*, the definite article *the* indicates that I am presupposing you know which snake I am talking about. If you do not, you would probably come back with a request for additional information to help you identify the referent of *the snake*, such as *which snake?* On the other hand, if I believed that you had no knowledge of the particular snake, I would have initiated the conversation with the NP indicated as indefinite: *Hey, Egbert just bought a snake!* Note that definite NPs are always presupposed, but the converse does not hold: if the referent of an NP is known to the speaker, but not the addressee, it could be claimed to be presupposed in the universe of the ongoing discourse, but is clearly *in*definite.

Topic NPs as typically presupposed and given are closely correlated with definiteness, and because pivot selection is largely equivalent to topic choice in English, indefinite pivots in English (and many other languages) are sometimes impossible:

(121)   (a) TOPIC   ‖   COMMENT
              John   ‖   has a new camera

              PRESUPPOSED ‖ FOCUS

              TOPIC   ‖ COMMENT
         (b) The/*a new camera ‖ is ‖ John's
              PRESUPPOSED   ‖   ‖ FOCUS

The variant with the indefinite NP as pivot is ungrammatical.

Personal pronouns like *I*, *you* and *she*, and proper nouns like *Hasan*, *Egbert* and *Australia*, are usually definite; they refer to entities which the speaker presupposes the addressee can uniquely identify. There can be exceptions, however; for example, when two people are looking over a map and one says to the other:

(122)   Hey, Fred, I found an Athens in Georgia!

Here, *Athens*, although a proper noun, is indefinite, because the speaker assumes the listener does not have this particular Athens in mind.

The distinction between definiteness and indefiniteness interacts with the previous two distinctions discussed. Definite nouns are presupposed and hence typically given information. They can be new, however, as in:

(123)   Hey, get up! The sun has already come up!

Here *the sun* is new information being introduced for the first time. Yet the speaker treats it as definite for he assumes the listener can uniquely identify the referent; there is only one sun in the real world. Definiteness interacts with referentiality to form all four possible combinations: (i) definite and referential – *I'm looking for the snake* (our pet diamond python); (ii) indefinite and referential – *I'm looking for a snake* (my diamond python just escaped from its cage); (iii) definite and nonreferential – *I'm looking for the friendliest snake I can find* (I haven't got it yet, but you and I know what kind it would have to be); and (iv) indefinite and nonreferential: *I'm looking for a snake to buy* (said to a pet-shop owner).

The final distinction in the information status of NPs to be considered here is that between generic and specific. This distinction indicates whether the NP refers to the entire class of its potential referents (generic) or a specific one:

(124)   (a) Snakes are easy pets to care for
          (b) The snake is my favourite among my pets, but it keeps escaping

Snakes in (124a) is generic; the statement is meant to cover all animals classed as snakes. The NP *the snake* in (124b) is specific, as the sentence only applies to a particular snake, my pet snake. Generic NPs can be definite or indefinite in English:

(125)   (a) The snake is an easy pet to care for
          (b) A snake is an easy pet to care for
          (c) Snakes are easy pets to care for

We can summarize the discourse-based distinctions of information status of NPs as:
 (i) Referent of NP is already established in discourse
            Yes = given/presupposed information
            No = new information
 (ii) NP refers to particular entity
            Yes = referential
            No = nonreferential (further, if NP is to be taken as referring to entire class of possible referents, then = generic)

(iii) Speaker presupposes addressee can uniquely identify referent
     Yes = definite
     No = indefinite

3.3   *The animacy hierarchy*

The distinctions of information status of NPs looked at in the last section were all established by the discourse context. I will now investigate how *inherent* properties of the referents of NPs determine their information status – in particular, the importance of being one of the participants in the immediate speech act, the speaker or addressee. Speaker and addressee are typically marked by the first and second person pronouns, *I* and *you* in English, although pronominal affixes on verbs are another common realization in many languages, e.g. Yimas. Traditionally, pronouns were claimed to stand for nouns, hence their name, 'for-noun', but this is clearly inaccurate in the case of first and second person pronouns for there is no possible NP which they could plausibly be argued to replace. The referents of *I* and *you* are not constant, but shift during the course of a verbal interaction, depending on who is doing the talking and who is spoken to: *ego* 'I' when said by Julius Caesar did not have the same referent as when said by Brutus. This constant shifting of referents of first and second person forms is a pervasive fact about human language.

The third person pronouns, on the other hand, do fit the traditional definition of pronouns as forms which stand in for nouns. English *he*, *she*, *it*, *they* are fundamentally different from *I*, *we*, *you* because they do not have the participant of the immediate speech act as referents; rather, their referents are restricted to non-participants in the speech act. They may refer to any referent other than the presently speaking person or the person(s) being spoken to. Third person is in essence a non-person in the context of the speech event.

All languages regard the speech act participants of first and second person as more salient than the non-person of the third. This is a fundamental principle that shows up in the grammatical structures of many languages (Silverstein (1976)). Beyond this, there is some variation as to the relative ranking of first and second person. In some languages they seem to have equal salience. In others, the speaker outranks the addressee, giving a hierarchy of speaker > addressee > third person. Many languages exhibit further elaborations of this person salience hierarchy, distinguishing different kinds of third persons, usually along the lines of humanness or animacy (Silverstein (1976)), so that this is perhaps better described as an animacy hierarchy. Tlahuitoltepec Mixe, a language of the Mixe-Zoque family of Mexico (Lyon (1967)), illustrates this well; its animacy hierarchy is speaker > addressee > human proper > human common > other animate > inanimate.

Mixe verbs have verbal pronominal affixes which reference the highest participant along this hierarchy, regardless of whether it is [+A] or [−A]. The actual affixes used indicate whether the [+A] participant is higher or the [−A] is:

	[+A] HIGHER	[−A] HIGHER
speaker	n-	š-
addressee	s-	m- . . . -ə
third person	t-	y- ∼ -y- . . . -ə

(126)   human > animate

    (a) tə   paat ha həyuhk t-wopy
        PAST Peter ART animal 3.[+A]HIGHER-hit
        'Peter hit the animal'

    (b) tə   paat ha həyuhk w[-y-]opy-ə
        PAST Peter ART animal 3.[−A]HIGHER-hit-[−A]HIGHER
        'The animal hit Peter'

Note that the word order of these two sentences is the same, but the meanings are exactly opposite. This is due to the pronominal affixation on the verb. We have two [−oblique] arguments, a third person proper NP *paat* 'Peter' and a third person animate NP *ha həyuhk*. By the Mixe animacy hierarchy given above, *paat* 'Peter' outranks *ha həyuhk* 'the animal' and so precedes it in linear order in the clause. When *paat*, the higher-ranked participant, functions as [+A], as in (126a), the verb carries the prefix *t-*, but when it is [−A], the verb takes the circumfix-infix *-y-* . . . *-ə*, as in (126b). Similar considerations apply to (127a, b):

(127)   proper humans > common humans

    (a) tə   paat ha hɔɔˀy t-wopy
        PAST Peter ART person 3.[+A]HIGHER
        'Peter hit the person'

    (b) tə   paat ha hɔɔˀy w[-y-]opy-ə
        PAST Peter ART person 3.[−A]HIGHER-hit-[−A]HIGHER
        'The person hit Peter'

In (127) we again find the proper human NP *paat* 'Peter', but this time with the common human NP *ha hɔɔˀy* 'the person'. In the Mixe animacy hierarchy, *paat* 'Peter' still outranks *ha hɔɔˀy* 'the person' because it is proper as opposed to common, and so must precede the latter in linear order. When *paat* 'Peter' is [+A], the third person [+A] higher prefix *t-* occurs in the verb, but when it is [−A], we find the circumfix-infix *-y-* . . . *-ə* ('3.[−A]HIGHER').

The other rankings in the Mixe animacy hierarchy are supported by the following examples; note that, in every case, the higher-ranked participant

occurs first in linear order and is referenced by pronominal affixation to the verb:

(128)   second > third
        (a) tə   mehc ha   hɔɔˀy s-wopy
            PAST 2SG   ART person 2.[+A]HIGHER-hit
            'You hit the person'

        (b) tə   mehc ha   hɔɔˀy m-wopy-ə
            PAST 2SG   ART person 2.[−A]HIGHER-hit-[−A]HIGHER
            'The person hit you'

(129)   first > third
        (a) tə   əhc ha   hɔɔˀy n-wopy
            PAST 1SG ART person 1.[+A]HIGHER-hit
            'I hit the person'

        (b) tə   əhc ha   hɔɔˀy š-wopy
            PAST 1SG ART person 1.[−A]HIGHER-hit
            'The person hit me'

(130)   first > second
        (a) tə   əhc mehc n-coky
            PAST 1SG 2SG   1.[+A]HIGHER-want
            'I wanted you'

        (b) tə   əhc mehc š-coky
            PAST 1SG 2SG   1.[−A]HIGHER-want
            'You wanted me'

   This Mixe pattern, in which the inherent salience of NPs along an animacy or person hierarchy of relative salience affects the information structure packaging of the [−oblique] NPs in the clause, is common among the languages of the world and is usually called a 'direct–inverse system' (see articles in Givón (1994)), 'direct' being the construction which marks the [+A] as being higher on the animacy hierarchy, and 'inverse', the forms indicating the [−A] to be higher. Direct–inverse systems are extremely common in Amerindian languages. Besides Mixe-Zoque, language families which exhibit it include Algonkian (Wolfart (1973); Dahlstrom (1986)), Kiowa-Tanoan (Kroskrity (1985); Klaiman (1991)), Athabaskan (Witherspoon (1980)), Cariban (Gildea (1994)), Tupi-Guarani (Payne (1994)), Araucanian (J. Grimes (1985)) and, more controversially, Wakashan (Whistler (1985); Emanatian 1988)) and Salishan (Jelinek and Demers (1983)). They are, however, not restricted to Amerindian languages, also being distinctive of Tibeto-Burman languages (DeLancey (1981)), like Chepang (C. Thompson (1990)).

3.4   *Topics, pivots, and prominence*

We have seen that topics and pivots share a number of properties. Not only is the pivot the typical choice for topic in languages like English, but, cross-linguistically, the characteristic properties of pivots and topics show a great deal of overlap. But it is also clear from languages like English that we cannot completely equate pivots and topics, for sentences exist in which they are distinct:

(131)

	TOPIC	PIVOT	
(a)	Soukous,	I	heard it everywhere in the Congo
(b)	Thai cooking,	he	can't get enough of it
(c)	Fred,	I	just saw him in the computer room

Let us call these non-pivot clause-initial topic NPs *external topics*. Such constructions are pervasive in the world's languages (for Mayan examples, see Aissen (1992)). In sentences like those of (131), both the external topic and the pivot display features of information status typical of topics, although external topics are commonly (but not always) contrastive or non-presupposed, while pivots are typically continuing, presupposed topics (of course, pivots need not be topics at all in English: they can be focus NPs: *Who saw Ian?*). External topics are not constituents of the clause, but are *external* to it, adjoined to the clause as a whole (see section 5 and chapter 3 by Andrews). And again unlike pivots, they are not restricted to [−oblique] arguments. Indeed, external topics seem the least marked with [+oblique] NPs, more so with [−oblique], [−pivot] NPs and most marked with [−oblique], [+pivot] NPs:

(132) (a) In the morning I finished the article
      (b) In Sydney there is always lots to do
      (c) With this new computer I'll be able to do a lot more
      (d) Egbert, I couldn't find
      (e) That movie, I just couldn't watch
      (f) Sam, he's going to die tomorrow
      (g) Soukous, it's the greatest African twentieth-century gift to
          civilization

Note the prosodic and syntactic differences among the sentences in (132). When [+oblique] NPs function as external topic (132a–c), there is no pause between them and the following clause. When [−oblique] [−pivot] NPs appear

as external topics, as in (132d, e) the pause is noticeably longer, but when [−oblique] [+pivot] NPs are external topics (132f, g), a longer pause is by itself not enough. A resumptive NP such as a pronoun must appear in the pivot position, where the external topic NP would be expected to be found if not topicalized.

Having established the contrast between external topics and pivots, there still remains the problem of explaining the correlations we have seen between topics and pivots. The answer seems to lie in the correlation of topics and pivots with the discourse-given features of information status of NPs, in particular the concept of presupposed, established information or givenness. NPs which refer to presupposed given referents, hence known and established, tend to be attenuated in formal realization. There is some cross-linguistic variation in this area, but the basic parameters of choice seem to be zero realization, i.e. control or zero anaphor; pronominal realization, as either a free pronoun or a bound verbal pronominal affix, or a full noun-headed NP. So, sentences like *Egbert saw Mary at the party and ϕ asked her to ϕ go out with him* illustrate exactly this pattern. In the first clause, Egbert and Mary are first introduced, and, not being established or given prior to this clause, they are realized by full NPs. But by the second conjoined clause and the infinitive complement, they are fully given and so can be realized by ellipsis ϕ or controlled ϕ or by pronouns *her* and *him*. The more highly given or presupposed an NP is, the more extreme its degree of attenuation, as a rule of thumb. Topics are prototypically realized in attenuated form due to their high degree of givenness. But pivots are also prototypically realized in attenuated form; in one construction after another in section 2.2, in English, Dyirbal, and Tagalog, pivots are controlled NPs, realized as ϕ. But control structures are nothing more than highly syntacticized construals of givenness in which the referents of the controlled NP must be determined within the confines of the immediate sentence: *John asked Mary [to ϕ go out with him]*. The referent of ϕ must be *Mary*, but this is not necessarily true of *him*; it could be some other man, given the right context: *Egbert really liked Mary but was very shy. So his friend John asked Mary to go out with him*. So, while the referent of the pivot-controlled ϕ is necessarily bound to the immediate sentence (indeed the immediate main clause), this is not true of the [+oblique] free pronominal. Thus, what topics and pivots share is their prototypical high prominence on a scale of givenness. This is not necessary – pivots can be focus, and topics, new information – but it is prototypical. As a result, both pivots and topics tend toward attenuated realization, and to be equated: an English pivot is normally topic, and if you want to alter topic choice, switch pivot choice (see examples (106–9)). In a very real sense, pivots are syntacticized topics, choices which are determined by syntactic considerations, like constituent structure, verbal agreement, lexical argument structure, control, etc. Being a required syntactic

notion, pivots have a certain degree of autonomy from a pragmatic notion like topic, but are nonetheless closely correlated for all the reasons we have been discussing.

Of course, many, if not most, languages, lack pivots, and it is of interest to find out how the principles of topicality prominence work out in these languages. It turns out, as we might expect, that the prominence relations that inform argument structure, [+A] > [−A] > [ ], also determine relative likelihood of topicality. The [+A] argument is simply the most topic-worthy NP, and its position in this regard is no doubt buttressed by its prototypical animacy and often high position on the animacy hierarchy. In an asymmetrical pivot nominative–accusative language like English, the unmarked choice for pivot, and hence topic, is the most prominent argument in the argument structure, the [+A] of a transitive verb (see (110)). Again the [+A] is the most topic-worthy NP. But, perhaps most surprisingly, Du Bois (1987) has shown that some features of this type of information structuring are even found in ergative–absolutive languages whose pivot is arguably the [−A] of a transitive verb. In the Mayan language Sacapultec, of the two arguments of a transitive verb, the [+A] and [−A], the [+A] argument has a much higher frequency as an attenuated bound pronominal realization (87 per cent of the time as opposed to 53 per cent for [−A] NPs). Further, the [+A] is normally given information. In fact there is an informal constraint to avoid non-presupposed [+A] NPs, while [−A] are often new information (only 3 per cent of [+A] NPs are new information, while 25 per cent of [−A] NPs are – a ratio of 8 to 1!). In terms of the principles outlined here, [+A] arguments are more topic-worthy than [−A] arguments in Sacapultec, in spite of its [−A] pivot choice. Similar findings have been reported for other ergative–absolutive languages like Chamorro (Cooreman (1987)) and for symmetrical languages like Tagalog (Cooreman, Fox, and Givón (1984)). These findings suggest the following principle: cross-linguistically, the most prominent argument in the argument structure is the unmarked choice for topic; various syntactic options can alter this, e.g. pronominal agreement, different pivot selections via active or passive voice, external topics, etc., but it always remains the neutral option, regardless of the overall syntactic typology of the language. Only further work will demonstrate whether this strong claim holds up; I present it here as a working hypothesis.

4   On voice: clause-internal packaging options

4.1   *Passive constructions*

Passive is a lexical process of verbal derivation that affects the linking between the levels of argument structure and grammatical functions (see also chapter 6 by Keenan and Dryer). Essentially, passive blocks the linking of the [+A] argument

to a [−oblique] function. This blocking is accomplished with some derivational affix on the verb or verbal complex. Consider the following English example of an active–passive contrast:

(133)   (a) The boy threw the ball
         (b) The ball was thrown by the boy

The active clause is produced by the normal linkings between levels already established:
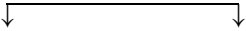
(134)   The boy threw the ball
              | |
           [+A]   [−A]
              | |
        [−oblique]   [−oblique]
              | |
         [+pivot]   [−pivot]

The form *be thrown* is the passive derived form of the lexeme *throw*. The effect of the passive derivation is to block the normal linking of [+A] to [−oblique] and, consequently, [+pivot]. Because all clauses must have pivots in English, the pivot then falls to the only other [−oblique] NP, the [−A] argument:
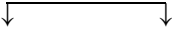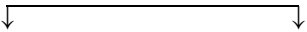
(135)   The ball was thrown by the boy
              | |
           [−A]   [+A]
              | |
        [−oblique]   [+oblique]
              |
         [+pivot]

   It is important to note that passive is a lexical process that affects the *linking* between argument structure and grammatical functions, *not* argument structure itself. In other words, the prominence relations among the arguments of the argument structure, [+A] > [−A] > [ ], may not be affected, so that [+A] arguments may still be able to participate in grammatical processes subject to their high prominence in argument structure. So, in Marathi (Joshi (1989)), [+A] arguments control reflexivization and are the target for control of nonfinite participial clauses:
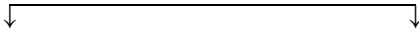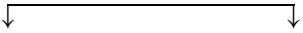
(136)   (a) jon-nii bil-la   aaplyaa gharaat maarle
             PN-ERG   PN-ACC   self's   house   hit
             'John$_i$ hit Bill$_j$ in self's$_{i/*j}$ house'

(b) [ɸ gharii jaa-uun] jon-nii bil-laa shaalet paathawle
home go-NFN PN-ERG PN-ACC school sent
'On ɸ_{i/*j} going home, John_i sent Bill_j to school'

The same grammatical properties still accrue to the [+A] argument in passive clauses, demonstrating that the [+A] argument does not lose its most prominent position in argument structure in these clauses:

(137) (a) bil-laa jon-kaduun aaplyaa gharaat maarle gele
PN-ACC PN-by   self's house hit   PASS
'Bill_i was hit by John_j in self's_{*i/j} house'

(b) [ɸ gharii jaa-uun] jon-kaduun bil-laa maarle gele
home go-NFN PN-by   PN-ACC hit   PASS
'On ɸ_{*i/j} going home, Bill_i was hit by John_j'

In English, however, the [+A] argument does cede these properties to the [−A] pivot NP:

(138) (a) Bill_i was given by John_j a picture of himself_{i/*j}

(b) While ɸ_{i/*j} watching the film, Egbert_i was tickled by Sam_j

The differences between Marathi and English seem explicable in terms of secondary language-specific effects of the passive lexical derivation. The universal effect is as we have defined it: blocking the linking of the [+A] argument to [−oblique]. For some languages like Marathi, this is as far as it goes; note, for example, that the Marathi [−A] argument even remains in accusative case in the passive! But, for other languages other effects occur. In English, pivots are required, so the only other [−oblique] NP, [−A] argument, must assume pivot status. Constructions sensitive to pivot status, like control, are then targeted by it. Further, the [+oblique] status of the [+A] argument in passive constructions requires that it cede its ability to bind reflexives, which is typically a property of [−oblique] NPs in English, not [+A] argument: *John_i gave Mary_j a picture of himself/herself*; *John_i gave a book to Mary_j about himself/*herself. John* and *Mary* are both [−oblique] NPs in the first example, so either may bind the reflexive. But in the second example, only *John* is [−oblique], so it is the only acceptable binder.

Passives are found in languages of all types: pivotless, asymmetrical nominative–accusative (i.e. English), asymmetrical ergative–absolute, and,

perhaps surprisingly, symmetrical. But in symmetrical languages there is, in fact, nothing preventing the removal of the [+A] argument to [+oblique] status, provided, of course, that it is not pivot, i.e. pivot choice is [−A]. Such patterns are found in some Philippine languages, such as Sama (Foley and Van Valin (1984)). The normal construction for a [−A] pivot choice is (139a); its passive equivalent is (139b):

(139)
(a) φ-b'lla   d'nda   kiyakan   kami
    pivot=[−A]-buy girl [+A] food [−A] 1PL.POSS
    'The girl bought our food'

(b) φ-b-*i*-'lla   uk   d'nda   kiyakan   kami
    pivot=[−A]-[PASS]-buy [+oblique] girl [+A] food [−A] 1PL.POSS
    'Our food was bought by the girl'

The passive formed is marked by the infix *-i-* and the [+oblique] status of the [+A] argument. The [−A] argument, of course, undergoes no change because it is already the pivot:

(140)   (a) normal ([−A] pivot)

                                          [+A]   [−A]
                                           | |
              φ-b'lla 'buy'   <x,   y>
                                           | |
                                      [−oblique]   [−oblique]
                                           | |
                                      [−pivot]   [+pivot]

         (b) passive

                                          [+A]   [−A]
                                           | |
              φ-b-i-'lla 'was bought'   <x,   y>
                                           | |
                                      [+oblique]   [−oblique]
                                                              |
                                                         [+pivot]

Passive constructions are also found in asymmetrical ergative–absolutive languages. As with symmetrical languages, they leave the pivot [−A] argument unaffected, but block the assignment of [−oblique] to the [+A] argument. Mam, a Mayan language of Guatemala, is one such language (England (1988)):

(141) (a) ma   ɸ-jaw   ky-tx'ee'ma-n   xiinaq   tzee'
           PAST   3SG.ABS-DIR   3PL.ERG-cut-DIR   man [+A]   tree [−A]
                  [−A]   [+A]
           'The men cut the tree'

       (b) ma   ɸ-tx'eem-at   tzee'   (ky-u'n xiinaq)
           PAST   3SG.ABS-cut-PASS   tree [−A] 3PL-by   man
                  [−A]
           'The tree was cut by the men'

Example (141a) is a fully transitive ergative–absolutive construction with pronominal affixes in the verbal complex for both [+A], *ky-* 3PL.ERG and [−A] as ɸ. Both [−oblique] full NPs occur following the verb with no prepositions. The passive in (141b) no longer has pronominal affixation for [+A]; as a [+oblique] NP due to passive, it is not eligible for this. Now the only [−oblique] NP is the [−A] argument *tzee'* 'tree', which occurs as above, an NP without preposition realized as a ɸ prefix on the passive, *-at* suffixed verb (other persons would have overt prefixes). Any [+oblique] [+A] necessarily co-occurs with the agentive preposition *-u'n* 'by', but could be freely omitted.

4.1.1   Foregrounding passives   Passive constructions can be divided into two types depending on their basic function. I adopt a functional typology of voice first proposed in Foley and Van Valin (1984). Foregrounding passives are those constructions whose function is to highlight the [−A] argument; typically, foregrounding passives make the [−A] argument topic. Formally, the [−A] argument always takes over the grammatical prerogatives of the [+A] argument, either as pivot in nominative–accusative languages like English, or as the sole [−oblique] NP and hence prominent argument in derived intransitive passive clauses in pivotless languages. In nominative–accusative pivot languages, foregrounding passives are normally required for the many constructions in which the pivot is not the [+A] argument:

(142)   (a) *Egbert wants [Mildred to take ɸ to the ball]

        (b) Egbert wants [to ɸ be taken by Mildred to the ball]

(143)   (a) *A python seems [Egbert to have bought ɸ]

        (b) A python seems [to ɸ have been bought by Egbert]

(144) (a) *Ian expects a python [Egbert to buy ɸ]

       (b) Ian expects a python [to ɸ be bought by Egbert]

(145) (a) *The guy [the bartender kissing ɸ] is my brother

       (b) The guy [ɸ being kissed by the bartender] is my brother

(146) (a) *Fred walked to the library and Bill met ɸ

       (b) Fred walked to the library and ɸ was met by Bill

In each of these cases the target of control or ellipsis is a [−A] argument of a transitive verb. The target of these processes is necessarily the pivot in English and other languages. Active sentences such as the (a) examples are therefore ungrammatical because they present the [+A] argument as pivot. Rather, fore-grounding passives are necessary, as in the (b) examples, which not only put the [+A] argument into a [+oblique] status, but also require the [−A] argument to assume the vacated properties of the [+A], in this case pivot status. The purpose of the foregrounding passives in (142)–(146), then, is to get the [−A] promoted in its syntactic status, to assume the grammatical properties of the pivot.

   Considerations of the animacy hierarchy also play a role in the func-tion of foregrounding passives. When the [−A] argument is of significantly higher rank on the animacy hierarchy than the [+A] argument, foreground-ing passives, which raise the syntactic prominence of the [−A], are often favoured:

(147) (a) I got run over by a bus
       (b) Egbert was stung by a bee!

They are strongly disfavoured in the case of the opposite situation:

(148) (a) ?A bee was killed by Egbert
       (b) ?A bus was run off the cliff by me

4.1.2 Backgrounding passives   The function of backgrounding passives is in a sense the core function of passive: to remove the [+A] argument from prominence in the clause. The degree of downgrading is a parameter that varies

quite widely from language to language. We have already seen one language, Marathi, in examples (136–7) in which the [+A] argument, although realized as a [+oblique] NP, as it must be in a passive construction, still controls a number of syntactic properties due to its being the most prominent argument in the argument structure. Other languages have constructions like the Marathi passive, but in which the now [+oblique] [+A] argument does cede these properties, and in fact is completely suppressed from the clause; for example Ulcha, a Tungusic language of Siberia:

(149)   ti   dūse-we   hōn-da   ta-wuri
         DEM tiger-ACC how-PTCL do-PASS
         'What's to be done about that tiger?'   Nichols (1979)

The verb *ta-* 'do' is lexically derived as passive and intransitive via the suffix *-wuri*. As a passive construction, the [+A] argument is blocked from linking to [−oblique]. In fact Ulcha, like many languages, applies a stronger restriction: not only must the [+A] argument not be [−oblique], it must be removed from the clause entirely, obviously prohibiting any control and binding properties such as are found in Marathi passives. Note, further, that, unlike in foregrounding passives, the [−A] remains in accusative case, unable to take on even the nominative case now vacated by the suppressed [+A] argument.

In other languages with backgrounding passives, the [−A] argument does assume the case marking or other morphological properties of the suppressed or now [+oblique] [+A] argument. This is because these are tied to the properties of argument structure: the [−A] argument, now the only [−oblique] argument in the argument structure of the now derived intransitive verb, is treated identically to the sole [−oblique] argument of an underived intransitive verb and will display whatever morphological properties are proper to that. Chickasaw (150) (Munro and Gordon (1982)) and Fijian (151) illustrate this pattern (in Chickasaw, passive is signalled by an infix *-hl-*):

(150)   (a) hattak-*at* ɸ-malili   V$_{INTR}$
             man-NOM   3SG[+A]-run
             'The man runs'

         (b) hattak-*at* ahi'   ɸ-ɸ-atahli   V$_{TRANS}$
             man-NOM   potato 3SG[+A]-3SG[−A]-prepare
             'The man prepares potatoes'

         (c) ahi'-*at*   ɸ-a<hl>taha   V$_{PASS}$
             potato-NOM 3SG[−A]-prepare <PASS>
             'The potatoes are prepared'

(151)   (a) e   lako na   gone V$_{INTR}$
            3SG go   ART child
            'The child goes'

        (b) au   kau-t-a   na   gone   V$_{TRANS}$
            1SG[+A] carry-TRANS-3SG[−A] ART child[−A]
            'I carried the child'

        (c) e   kau-ti   na   gone V$_{PASS}$
            3SG carry-PASS ART child
            'The child was carried'

Note that, in both Chickasaw and Fijian, the [−oblique] [−A] argument of a passive construction (the (c) examples) has the same grammatical properties as the sole [−oblique] argument of an underived intransitive verb (the (a) examples) and is different from when it functions as the [−A] argument of a transitive verb (the (b) examples). In Chickasaw, the [−A] argument fails to receive the -(a)t nominative case marker for the most prominent NP, if it is an argument of a transitive verb (150b); while as the sole [−oblique] argument of both the derived intransitive passive verb (150c) and an underived unergative intransitive verb (150a), it does so, for it is now the most prominent argument, there being no other [−oblique] NPs. In Fijian, the [−A] argument of a transitive verb is realized with suffixed pronominals, -a 3SG [−A] in (151b). However, the sole [−oblique] argument of underived intransitive verbs co-occurs with *proclitic* pronominals (151a). Note that the sole [−oblique] [−A] argument of a derived passive clause in Fijian has the same proclitic pronominals (151c). The evidence suggests that languages like Chickasaw and Fijian go much further than Marathi in the backgrounding of the [+A] argument: its [+oblique] status or complete suppression forces it to cede its prominent position in argument structure to the only [−oblique] NP, the [−A] argument. This is a language-specific side effect of passive in such languages, not a universal constraint, as languages like Marathi demonstrate. These passive constructions in Chickasaw and Fijian are clearly backgrounding passives because both languages are pivotless, and, further, the main function of the construction is to background or suppress the [+A], not topicalize or foreground the [−A].

Some European languages, notably the asymmetrical nominative–accusative pivot languages of the Germanic family, have an interesting intermediate passive construction reminiscent of Ulcha, in which the [−A] of the passive construction
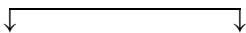
retains some of the properties of [−A] arguments of transitive verbs, but also takes on some of those of [−oblique] arguments of intransitive verbs. These are the impersonal passives illustrated from Dutch (Kirsner (1976); John Verhaar (personal communication)):

(152)   er   woorden   daar huizen gebouwd
          there became(PL) there houses built
          'There were houses built there'

Sentence (152) is formally a passive with the lexically derived passive form of the verb and the passive auxiliary *woorden* 'became'. The [+A] is completely suppressed in this example, but the [−A] argument does not take on the full properties of the pivot. Note that it remains in the post-auxiliary position of a [−A] argument of a transitive verb, rather than the pre-auxiliary position of the pivot. It does, however, trigger verb agreement (plural) in the manner of the pivot. The otherwise vacant pivot position is occupied by the 'dummy' pronoun *er* 'there'. This backgrounding passive construction is permissible whenever one wishes to background the [+A] argument, either presenting it as [+oblique] or suppressing it entirely; interestingly, it is not restricted to transitive verbs, but is available whenever the verb could occur with a [+A] argument, for example with unergative intransitive verbs:

(153)   (a) er   wordt   door de jongen gefloten
              there become(SG) by   the boys   whistled
              'There is whistling by the boys'

          (b) er   wordt   door de studenten gestaakt
              there become(SG) by   the students struck
              'There is a strike by the students'

These sentences are completely pivotless; neither the [+oblique] [+A] NP nor the dummy pivot holder can be the target for control:
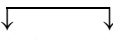
(154)   *er   werd   door de vrouwen gelachen en   φ huilden
          there became(SG) by   the women   laughed and   cried
          'There was laughing by the women and cried'

   Probably the most common usage of backgrounding passives is to present a resulting state which has affected the [−A] as a consequence of the action of a [+oblique] or, more usually, suppressed [+A] participant. English presents many examples of this type of backgrounding passive or *mediopassive*:

(155)   (a) The house was robbed yesterday
         (b) The roast was overcooked

Some languages have distinct passive constructions for the foregrounding and
mediopassive backgrounding function. This is particularly common among the
Bantu languages, for example, KiSwahili:

(156)   (a) chakula   ki-na-pik-w-a   na mama
             food.VII.SG VII.SG-PRES-cook-PASS-MOOD by mother
             'The food is being cooked by mother'

         (b) chakula   ki-na-pik-ik-a   (*na   mama)
             food.VII.SG VII.SG-PRES-cook-PASS-MOOD by   mother
             'The food is cooked'

The passive with -*w* is a foregrounding one. The [+A] NP can be overt, although
[+oblique], and the [−A] appears as the sole [−oblique] NP; further, the verb
remains dynamic aspectually. The form with -*ik* is a backgrounding passive:
again the [−A] argument assumes the status of the sole [−oblique] NP, but the
[+A] is necessarily suppressed and the verb is aspectually stative.

4.1.3   Summary   Table 7.1 summarizes our typology of passive construc-
tions. Foregrounding passives seem to be restricted to languages with pivots.
This makes sense in view of their function – to get the [−A] in a prominent
position for syntactic purposes, as pivot, a binder of anaphors or target for
ellipsis or control. But there is no a-priori reason why a language could not
have foregrounding passives for reasons other than access of [−A] arguments
to pivot status. It is certainly conceivable that a pivotless language could have
foregrounding passives, for example, to link the [−A] to topic function in a
main clause, but with none of the other diagnostic properties of pivots. We
would be justified in analysing this as a foregrounding passive in a pivotless
language. Backgrounding passives, as the unmarked type, show no such restric-
tions. They are found in languages of all types: pivotless (Ulcha), symmetrical
(Sama), asymmetrical nominative–accusative (English, KiSwahili) and asym-
metrical ergative–absolutive (Mam). The major parametric difference among
languages for backgrounding passives is the degree to which the [−A] assumes
the properties vacated by the now [+oblique] [+A]. At one extreme is Marathi
in which the [+oblique] [+A] argument seems to retain all or most of its promi-
nence in argument structure and grammatical properties tied to this. At the other
are languages in which the [+A] is completely suppressed and all grammatical

Table 7.1 *Summary of passive constructions*
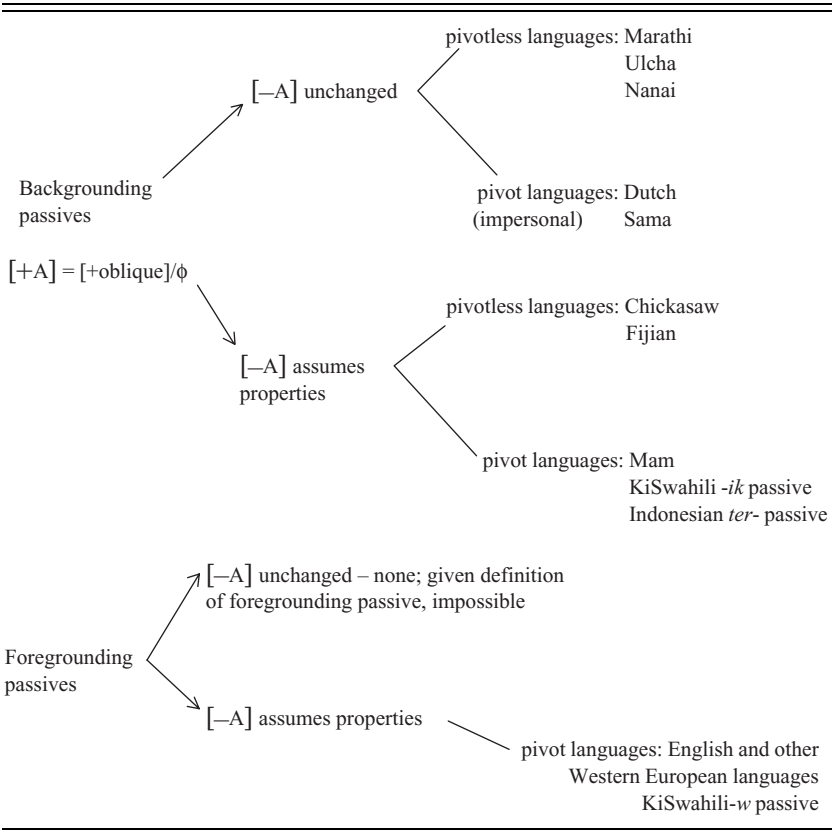
properties due to prominence now accrue to the sole [−oblique] NP, the [−A] argument. This parameter of variation, of course, is not available in foregrounding passives, the [−A] being required in all cases to assume the grammatical properties of prominence ceded by the [+oblique] [+A] argument. To do otherwise would be senseless, to fly in the face of what defines a foregrounding passive.

Even a cursory study of the world's languages reveals that not all have passive constructions, so an obvious question is: how do languages that lack a passive express the obviously useful function of backgrounding the [+A] argument? The usual way is to have a nonreferential NP filling the role of the [+A] argument. European languages like French and German have special impersonal pronouns to perform exactly this function:

(157) (a) on   parle   français içi
            one speaks French   here
            'French is spoken here'

       (b) Mann spricht Deutsch
            one   speaks   German
            'German is spoken'

Of course, French and German possess passive constructions, but some Amerindian languages which lack passive constructions have impersonal pronominal verbal affixes to fulfil this function, such as Caddo:

(158) diiwikkudah
       ɸ-yi-ˀawidakud-ah
       3SG[−A]-REALIS.IMPERSONAL[+A]-fire- PERF
       'One has fired him' = 'He's been fired'   Chafe (1990)

4.2   Antipassive constructions

Antipassives are those verbal lexical derivations which background the other [−oblique] argument of a transitive verb, the [−A] argument. An affix to the verb or verbal complex derives an intransitive antipassive verb form and the [−A] argument necessarily appears as [+oblique] or is suppressed entirely. As an intransitive form, the [+A] argument will now be indicated in the same morphological form as the sole [−oblique] argument of an intransitive verb. Antipassives are most prototypical of asymmetrical ergative–absolutive pivot languages. Consider these examples from the Mayan language Mam:

(159)
(a) o   chi-tzaj   t-tzyu-ˀn   Xwan   xiinaq
    PAST 3PL.ABS[−A]-DIR 3SG.ERG-grab-DIR John[+A]   man[−A]
    'John grabbed the men'

(b) o   ɸ-tzyuu-n   Xwan   ky-e   xiinaq
    PAST 3SG.ABS-grab-ANTIPASS John[+A] 3PL.POSS-[+oblique]   man[−A]
    'John grabbed the men'   England (1988)

Example (159b) is the intransitive antipassivized version of (159a). Example (159a) is a normal transitive clause: the verbal expression has pronominal affixes for both [−oblique] arguments, *chi-* (3PL.ABS[−A]) and *t-* (3SG.ERG[+A]), and both full NPs are [−oblique], with no case marking or adpositions. In the antipassive (159b), the [−A] argument *xiinaq* 'man' is now [+oblique] marked with the adposition *ky-e* (3PL.POSS-[+oblique]); note it fails to have a co-occurring pronominal affix on the verb. The [+A] argument *Xwan* 'John' is still [−oblique] and co-occurs with the zero third person pronominal absolutive affix, as the sole

[−oblique] argument of the derived intransitive verb. The verb occurs with the antipassive derivational suffix -*n* (ANTIPASS), marking this as a lexical derived intransitive verb.

Parallel to the passive, the antipassive may be analysed as a lexical derivational process that blocks the linking of the [−A] argument to [−oblique] function. In an asymmetrical ergative–absolutive pivot language like Mam, we have the following linkings for a normal transitive clause:

(160)
```
                    [+A]              [−A]
                     |                 |
    tzyu-'grab'     <x,                y>
                     |                 |
                [−oblique]        [−oblique]
                     |                 |
                 [−pivot]          [+pivot]
```
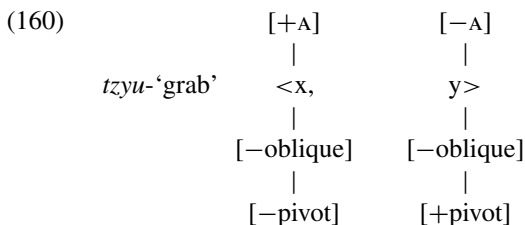
Antipassive blocks the linking of the [−A] argument to [−oblique] and, therefore, in an ergative–absolutive pivot language like Mam, to pivot status; by default the [+A] argument now becomes pivot:

(161)
```
                             [+A]              [−A]
                              |                 |
    tzyuu-n 'grab'-ANTIPASS  <x,                y>
                              |                 |
                         [−oblique]        [+oblique]
                              |
                          [+pivot]
```
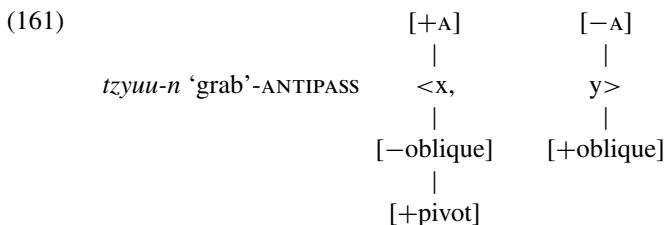
Unlike passives, which do not require the [−A] argument to take up the grammatical properties vacated by the [+oblique] [+A] argument (e.g. Marathi, Ulcha), antipassive constructions always present the [+A] argument as the prominent [−oblique] NP. This perhaps is due to the universal hierarchy of prominence, [+A] > [−A] > [ ], in which the [+A] NP universally has the highest degree of prominence in the argument structure.

4.2.1   Foregrounding antipassives   Foregrounding antipassives are those whose function is to remove the [−A] to [+oblique] in order to permit the [+A] to acquire a more prominent syntactic status, typically that of pivot, so that these constructions are diagnostic of asymmetrical ergative–absolutive pivot languages like Dyirbal or Mam.

Mam resembles English in that its pivot NP is syntactized as the target of a number of grammatical constructions, although the two languages differ, of course, in their unmarked choices for pivot, [+A] or [−A]:

(162)   English   Mam

V_TRANS   <x,   y>   V_TRANS   <x,   y>
             | | | |
          [+A]   [−A]   [+A]   [−A]
             | | | |
        [+pivot] [−pivot]   [−pivot] [+pivot]

In Mam [−A] arguments can be relativized freely, as befits their status as pivots:

(163)   ma-aˀ   ɸ-w-il-a
         TNS-EMPH   3SG.ABS-1SG.ERG-see-1SG
                           [−A]   [+A]

         tii-xiinaq   [x-ɸ-tzaj   ky-tzyuˀn   ɸ]
         INTENS-man   REL.TNS-3SG.ABS-DIR   3PL.ERG-grab-DIR
         'I saw the man whom they grabbed'

This is, however, not possible for [+A] arguments, which, as Mam is an asymmetrical ergative–absolutive language, are not pivots. In order for [+A] arguments to be relativized NPs, they must become pivots. To do this a foregrounding antipassive is used, which derives an intransitive unergative verb, with the [+A] argument as its sole [−oblique] NP and, by default, pivot of the clause:

(164)   ma   ɸ-w-il-a
         TNS 3SG.ABS-1SG.ERG-see-1SG
                 [−A]   [+A]

         tii-xiinaq   [x-ɸ-tzaj   tzyuu-n   ɸ
         INTENS-man   [REL.TNS-3SG.ABS-DIR   grab-ANTIPASS

         ky-e   xjaal]
         3PL.POSS-[+oblique] person
         'I saw the man who had grabbed the people'

Further, foregrounding antipassives are required in Mam whenever the [+A] argument is the target of control in nonfinite infinitival complements, just as with Dyirbal nonfinite purposive clauses (see section 2.3). Here the allomorph of the antipassive suffix is ɸ in contrast to the usual -n, but it is clear that

Table 7.2 *Foregrounding passives and antipassives*

	Asymmetrical NOM–ACC	Asymmetrical ERG–ABS
Basic pattern:	$[+A]/V_{TRANS} = [+pivot]$	$[-A]/V_{TRANS} = [+pivot]$
Derived pattern:	Passive: $[-A] = [+pivot]$	Antipassive: $[+A] = [+pivot]$

antipassive has applied to the complement because the $[-A]$ of a transitive verb appears in the usual $[+oblique]$ form of antipassives:

(165)
(a) o   chi   e²x   xjaal   [laq'oo-φ-l   φ t-ee]
    PAST   3PL.ABS   go   person   buy-ANTIPASS-NFN   3SG.POSS-[+oblique]
              [+A]
    'The people went to buy it'

(b) n-chi-ku²   teen   xjaal   [belaara-φ-l   φ
    PROG-3PL.ABS-DIR   begin   person   watch-ANTIPASS-NFN
              [+A]
    t-e   jun   weech
    3SG.POSS-[+oblique]   one   fox
    'The people began to watch the fox'

Patterns like (165) are fully expected in asymmetrical ergative–absolutive pivot languages like Mam (they have already been encountered in the typologically similar language Dyirbal in section 2.3). Because the $[-A]$ of a transitive verb is the unmarked pivot, it is the normal target for control. But for semantic reasons, it is typically the $[+A]$ argument which is the shared NP between main clause and complement in all languages (Dixon (1994)) and is therefore eligible for control. This conflict is resolved in asymmetrical ergative–absolutive languages like Mam through the use of a foregrounding antipassive, which derives an intransitive unergative verb from a lexically transitive one, through blocking the linking of the $[-A]$ argument to $[-oblique]$. As the derived verb is formally intransitive, the $[+A]$ argument is now the only $[-oblique]$ NP and therefore assumes the syntactic properties of pivot.

    Foregrounding passives and antipassives are mirror images of each other in languages of contrasting typologies, as is shown in table 7.2.

    Asymmetrical nominative–accusative languages like English and asymmetrical ergative–absolutive languages like Mam are mirror images of each other in their basic pivot choices between the two $[-oblique]$ NPs of a transitive verb, $[+A]$ and $[-A]$. In asymmetrical nominative–accusative languages, the choice is the $[+A]$ argument, but in asymmetrical ergative–absolutive languages it is

the [−A]. This can create problems in both kinds of languages when syntactic constructions sensitive to the pivot notion require the other [−oblique] NP to assume this function, as we have seen in many examples drawn from English, Dyirbal, or Mam in this and previous sections. In such situations, a construction is necessary to allow the other [−oblique] argument to assume the pivot function. In both types of languages, this is done by preventing the argument which functions as the unmarked pivot choice from linking to [−oblique], forcing it to appear as [+oblique]. There is now only one [−oblique] argument, so the verb is formally intransitive; in fact, the prototypical way this is done in these languages is through an affix to the transitive verb stem deriving an intransitive verb form. Because there is now only one [−oblique] argument, this will be the pivot and thereby will take on the syntactic properties proper to this function. In asymmetrical nominative–accusative languages, passive blocks the linking of the [+A] argument to [−oblique], so the [−A] argument necessarily becomes pivot, while in asymmetrical ergative–absolutive languages, antipassive does the same to the [−A] argument, so that the [+A] becomes pivot.

4.2.2   Backgrounding antipassives   Parallel to backgrounding passives, backgrounding antipassives represent the core function of antipassives; blocking the linking of the [−A] argument to [−oblique] status, without the necessary side effect of shifting pivot properties to the [+A] argument diagnostic of foregrounding antipassives. Consequently, backgrounding antipassives are much more common than foregrounding ones, being found in languages of more diverse typologies than foregrounding antipassives; in particular, they are quite wide-spread in pivotless languages. Because the [−A] argument in a backgrounding antipassive construction, if present at all, is necessarily [+oblique], the clause is formally intransitive. If the language distinguishes morphologically transitive from intransitive verbs, the verb of the clause will be intransitive, and, further, the sole [−oblique] argument of the derived antipassivized verb, the [+A] NP, will exhibit whatever grammatical properties are proper to the sole [−oblique] argument of intransitive verbs. Unlike backgrounding passives in languages like Marathi, Ulcha, or Dutch, which exhibit transitive properties, e.g. the [−A] argument remains in accusative case, this is not possible for backgrounding antipassives – they are necessarily formally intransitive: the ergatively case-marked [+A] NP of a transitive verb must, in the corresponding antipassive, assume the case proper to the sole [−oblique] argument of an intransitive verb. It would seem that the notion of transitivity is defined cross-linguistically in terms of the presence of the [−oblique] [−A] argument. Because backgrounding passives block the linking of the [+A] argument to [−oblique], they can leave the [−A] argument unaffected, so the derived passive clause can remain formally transitive. This is not possible for antipassives: blocking the

linking of the [−A] argument to [−oblique] status by definition makes the clause intransitive.

The most prototypical and wide-spread use of the backgrounding antipassive is for the complete suppression of the [−A] argument. Not only must it not be [−oblique], it cannot be overtly mentioned at all:

(166)   Bandjalang of Australia
(a) mala-yu   ḍa·ḍam-bu   mala   bulan   ḍa-ila
    DEM-ERG   child-ERG   DEM.ABS   meat.ABS   eat-PRES
                    [+A]   [−A]
    'That child is eating meat'

(b) mala   ḍaḍam   ḍa-le-ila
    DEM.ABS   child.ABS   eat-ANTIPASS-PRES
                    [+A]
    'That child is eating'   Crowley (1978)

The (a) example is fully transitive with two [−oblique] NPs, the [+A] and [−A] arguments. The (b) example is the corresponding backgrounding antipassive. The clause is formally intransitive, and the [−A] argument is fully suppressed.

In other cases, the now [+oblique] [−A] argument of an antipassive construction need not be fully suppressed. Its interpretation is, however, affected by its change in status and there are at least two ways in which this manifests itself. The notion of being totally affected by an action or undergoing a change in state is prototypically associated with [−A] arguments. In many languages, the effect of backgrounding antipassives is to strip this meaning from the [−A] argument; when [+oblique] in an antipassive construction, the [−A] argument is interpreted as partially affected in contrast to its normal interpretation in a full transitive clause:

(167)   Kabardian
(a) fie-m   q'ʷɨpŝfire-r jedzaq'e
    dog-ERG   bone-ABS   bite
        | |
      [+A]   [−A]
    'The dog bites the bone [through to the marrow]'

(b) fie-r   q'ʷɨpŝfire-m je-*w*-dsaq'e
    dog-ABS   bone-INSTR   [ANTIPASS]-bite
        | |
      [+A]   [−A]
        | |
    [−oblique] [+oblique]
    'The dog gnaws at the bone'   Catford (1976)

(168) Chamorro
    (a) un-patek   i   ga'lagu
          2SG.ERG-kick DET dog.ABS
              | |
          [+A]   [−A]
          'You kicked the dog'

    (b) mam-(p)atek   hao   gi   ga'lagu
          ANTIPASS-kick 2SG.ABS [+oblique] dog
                     | |
                [+A]   [−A]
                 | |
             [−oblique]   [+oblique]
       'You kicked at the dog'   Cooreman (1988)

In each language, the [−oblique] [−A] NP in the transitive (a) constructions
is fully affected (the bone is bitten through and the dog is actually kicked),
but in the antipassive (b) examples this does not hold (the bone is only being
gnawed at the surface and the dog may have escaped the kick). Note that
though the English translations show the same type of syntactic alternations
and corresponding meaning shifts, they do not illustrate true antipassives, as
there is no requisite derivational morphology for antipassivization in the verb
or verbal complex.

   In other cases, the effects of the backgrounding antipassive may be more
pragmatic than semantic: most commonly, the now [+oblique] [−A] argument
loses its ability to be referential or definite:

(169) Chukchee of Siberia
      (a) ˀaaček-a   kimitˀ-ən ne-nlˀetet-ən
          youth-ERG load-ABS 3PL.ERG-carry-3SG.ABS.AOR
          'The young men carried away the load'
      (b) ˀaaček-ət   ine-nlˀetet-gˀet   kimitˀ-e
          youth-ABS.PL. ANTIPASS-carry-3PL.ABS.AOR load-INSTR
          'The youths carried away a load'
                     Kozinsky, Nedjalkov, and Polinskaja (1988)

   Some languages, notably Mayan languages like Mam (England 1983b), allow
nonreferential [−A] arguments in backgrounding antipassive constructions to
be incorporated into the verbal complex:

(170) (a) n-ɸ-tx'aa-n   t-q'ool   poon
            PROG-3SG.ABS-chew-ANTIPASS 3SG.POSS-sap copal
            'She was copal sap-chewing'

       (b) ma   ɸ-b'iincha-n   qa-jaa
            PAST 3SG.ABS-make-ANTIPASS PL-house
            'He house-constructed'

This noun incorporation is licensed through the antipassive derivation: blocking
the linking of [−A] to [−oblique] allows it to be incorporated into the verbal
complex as a kind of [+oblique] adjunct with the usual generic nonreferential
meaning of incorporated nouns. No constituents can intervene between the
verbal complex and the incorporated [−A] NP; this can be the basis for a claim
that this is a noun incorporation construction.

However, noun incorporation need not involve antipassivization. Many lan-
guages have productive noun incorporation for generic [−A] NPs with no trace
of antipassive constructions, for example in what can be seen as noun incorpo-
ration in Kusaiean of Micronesia:

(171) (a) nga ɔl-læ   nuknuk ɛ
            1SG wash-PERF clothes DEF
            'I washed the clothes'

       (b) nga owo   nuknuk læ
            1SG wash clothes PERF
            'I clothes-washed'

       (c) nga owo   læ
            1SG wash PERF
            'I washed'   Sugita (1973)

The verb 'wash' occurs in two forms, a transitive stem *ɔl* (171a) and an intran-
sitive *owo* (171c). Example (171a) is a normal transitive clause, with the verb
suffixed with the aspect clitic *-læ* (PERF) and then followed by the [−oblique]
[−A] argument which is the definite *nuknuk ɛ* 'the clothes'. Example (171b) is
the corresponding form with noun incorporation. The [−A] argument is neces-
sarily generic and cannot co-occur with the definite determiner ɛ. The incor-
porated [−A] is no longer [−oblique], so the clause is formally intransitive;
note the verb form is the intransitive *owo* (compare (171c)). Further, the [−A]
argument, no longer being a full [−oblique] argument is incorporated into the
verbal complex like an adjunct, so that the perfective clitic *-læ* follows it. Note
that (171b) is not an antipassive: while the [+A] argument has failed to link
to [−oblique], there is no lexical derivation on the verb, no overt antipassive
suffix, licensing this. Noun incorporation, while related in its effects on the

[−A] argument, must be distinguished from antipassivization. This point is further emphasized by the fact that there are languages like Chukchee which have both antipassives and noun incorporation, and they are formally quite distinct. The Chukchee antipassive was illustrated in (169); note it makes use of an overt antipassive affix *ine-*. No such affix is found in Chukchee noun incorporation:

(172)   (a) tumg-e   n-antəwat-ən   kupre-n
             friend-ERG   3PL.ERG-set-3SG.ABS.AOR   net-ABS
             'The friends set the net'

        (b) tumg-ət   kupr-antəwat-gʔat
             friend-ABS.PL   net-set-3PL.ABS.AOR
             'The friends were net-setting'   Comrie (1978)

Here again, noun incorporation simply involves incorporating the nonreferential [−A] argument into the formally intransitive verb with no other needed derivational affixation; this is not antipassivization.

    Another construction that needs to be distinguished from antipassives is the use of impersonal pronominal forms for [−A] arguments, rather like impersonal [+A] forms such as French *on*. Such impersonal pronominals for [−A] arguments are not uncommon in Amerindian languages like Caddo:

(173)   nayt-ya-ʔiyah-hah
        he.who-REALIS.IMPERSONAL[−A]-catch-HABIT
        'he who catches one' = 'a policeman'   Chafe (1990)

Again these are not true antipassives: the verb remains transitive and the [−A], while indefinite and impersonal, remains [−oblique], as witnessed by its being realized as a pronominal verbal affix, a property restricted to [−oblique] arguments.

    Table 7.3 summarizes our typology of antipassive constructions.

4.3   Applicative constructions

In many ways applicatives are mirror images of the passive and antipassive constructions we have been discussing in the previous sections. Whereas passives and antipassives block erstwhile [−oblique] [+A] and [−A] arguments from being assigned [−oblique] status, relegating them to [+oblique], applicative constructions do the opposite: they take erstwhile [+oblique] NPs like recipients, benefactives, instruments, locationals or comitatives and make them [−oblique] [−A] arguments of the verb. Applicatives are found

Table 7.3 *Summary of antipassive constructions*
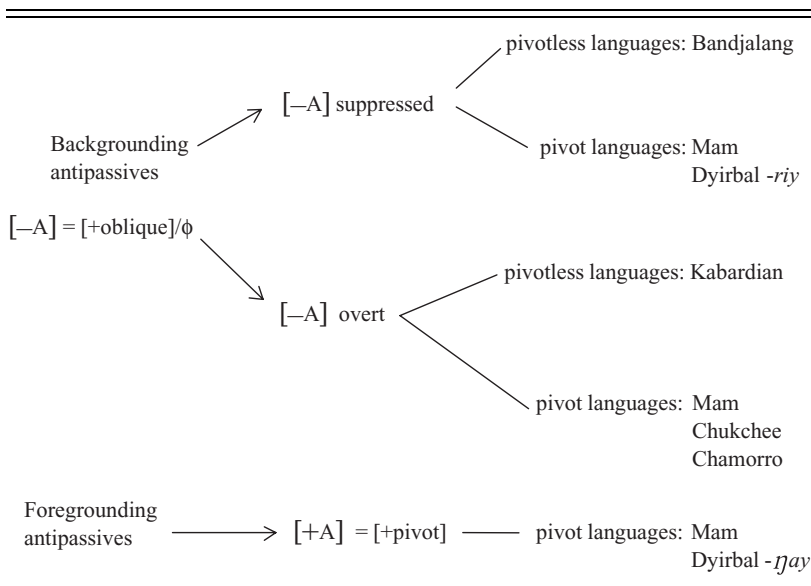
in languages of all typologies, pivotless, symmetrical, and asymmetrical, of
either the nominative–accusative or ergative–absolutive variety. Morphologi-
cally elaborate pivotless languages like Yimas are often especially rich in
applicative derivations. Like passive and antipassive, applicative formation
is a lexical process of derivation, marked by affixation to the verb or ver-
bal complex, but instead of denuding a verb of a [−oblique] argument, it
adds another one not specifically subcategorized by the verb. This is easiest
to see with an illustration; consider (174) with the Yimas intransitive verb
wa- 'go':

(174)   Yakayapan   na-nampan   pu-na-wa-n
          PN   3SG-toward   3PL[+A]-DEF-go-PRES
          'They are going toward Yakayapan'

The intransitive verb *wa-* 'go' has a sole [−oblique] argument, a [+A], realized
through the pronominal prefix *pu-* 3PL s. The clause also has a [+oblique]
adjunct, Yakayapan, marked as oblique by the postposition *nampan* 'toward'.
There is an alternative way to express the conceptual event described by (174),
and this is through the use of the allative applicative prefix *ira-* 'toward'. This

is a process of lexical derivation by which this prefix is added to the intransitive verb *wa-* 'go' to derive the now transitive verb *ira-wa-* 'go toward':

(175)   na-mpu-na-*ira*-wa-n   Yakayapan
        3SG[−A]-3PL[+A]-PROG-ALL-go-PRES   PN
        'They are going toward Yakayapan'

That the derived verb *ira-wa-* 'go toward' is now fully transitive is demonstrated by the presence of two pronominal prefixes on the verb: *na-* (3SG) for the [−A] argument and *mpu-* (3PL) for the [+A]. The participant *Yakayapan* which was [+oblique] in (174), governed by the postposition *nampan*, is now [−oblique] and realized as a verbal pronominal affix. This is the basis of applicative constructions: [+oblique] participants become [−oblique] and [−A] through a process of lexical derivation of the verb.

Applicative formation is also available to transitive verbs, in Yimas and other languages. Example (176a) is a sentence with a transitive verb and a [+oblique] adjunct; (176b) is its applicative counterpart:

(176)   (a) mpu-nampan   ŋarwa   wa-n-wampak-ɲcut
             3PL-toward   penis.IX.SG IX.SG[−A]-3SG[+A]-throw-REM.PAST
             'He sent his penis to them'

        (b) ŋarwa   wa-n-*ira*-wampak-ɲcuk-mpun
             penis.IX.SG   IX.SG[−A]-3SG[+A]-ALL-throw-REM.PAST-3PL[−A]
             'He sent them his penis'

The derived applicative verb in (176b) *ira-wampaki-* 'throw/send something toward someone' is a ditransitive verb, derived from the transitive verb root *wampaki-* 'throw/send something', as in (176a), plus the allative applicative prefix *ira-*, which adds the meaning component 'toward someone'. Note that, while the verb in (176a) is formally transitive, with two pronominal affixes, *wa-* (IX.SG[−A]) and *n-*(3SG[+A]), that of (176b) is ditransitive, having three pronominal affixes, realizing two [−A] participants, one with the absolute prefix *wa-*(IX.SG[−A]) and the other with the dative suffix *-mpun* (3PL[−A]), the latter corresponding to the [+oblique] participant in (176a). A ditransitive verb derived by applicative formation is indistinguishable in grammatical behaviour from a basic underived ditransitive verb like *ŋa-* 'give'. This process of derivation illustrated in (176) can be represented as (177):

(177)
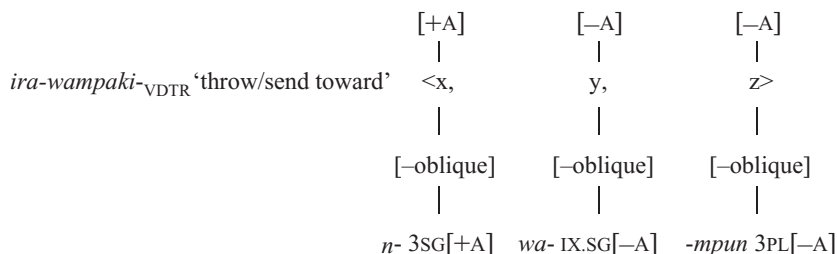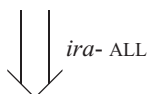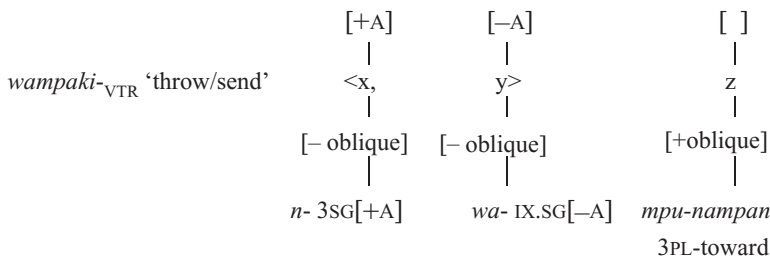
Applicative constructions in asymmetrical ergative–absolutive languages have additional complications. This is due to the one pivot per clause constraint. Because the unmarked pivot choice for asymmetrical ergative–absolutive languages is the [−A] argument, the one pivot per clause constraint prohibits multiple [−A] arguments in these languages. We have already seen in section 2.3 that these languages proscribe underived basic ditransitive verbs, so that verbs like 'give' in these languages are formally transitive. The same holds for verbs derived by applicative formation. They cannot be formally ditransitive. So when an applicative derivation does operate on a transitive verb root in these languages, it also has the effect of a backgrounding passive, so that the [−A] of the underived transitive verb appears as [+oblique], and the derived applicative verb is only formally transitive. Consider these examples from Dyirbal:

(178) (a) balan   ḍugumbil   baŋgul   yaṛa-ŋgu
            DET.ABS   woman.ABS   DET.ERG   man-ERG
            baŋgul   yugu-ŋgu   balga-n
            DET.INSTR   stick-INSTR   hit-TNS
            'The man is hitting the woman with the stick'

(b) bala   yugu   baŋgul yaṛa-ŋgu balgal-ma-n
    DET.ABS   stick.ABS   DET.ERG man-ERG hit-APPLIC-TNS
    bagun   ḍugumbil-gu
    DET.DAT   woman-DAT
    'The man is hitting the stick against the woman'

Dixon (1972)

Example (178a) is a normal transitive construction with the transitive verb *balgal-* 'hit' with two [−oblique] arguments, the [+A] *yaṛa-ŋgu* (man-ERG) 'man' and the [−A] *ḍugumbil* 'woman'. As Dyirbal is an asymmetrical ergative–absolutive pivot language, the pivot is the [−A], which has absolutive case, the case of [+pivot]. The clause also contains a [+oblique] instrument participant NP, *yugu-ŋgu* (stick-INSTR) 'with the stick' (the ergative and instrument case forms are homophonous, but Dixon (1972) presents convincing arguments that they are distinct; ergative is the case of [−oblique] NPs, while instrumental occurs with [+oblique]). Example (178b) is the corresponding applicative version of (178a), in which an applicative verb form *balgal-ma-* 'hit with' is derived. This derived form presents the former [+oblique] instrument *yugu* 'stick' now as [−oblique] and [−A]. As [−A] it necessarily assumes pivot status and appears in absolutive case. Because of the one pivot per clause constraint, the [−A] of (178a) *ḍugumbil* 'woman' can no longer be pivot and so, to prevent this, the applicativization also has the effect of a backgrounding antipassive, blocking the linking of the [−A] participant *ḍugumbil* 'woman' to [−oblique], so that it necessarily appears as [+oblique], in (178b) in the dative case, the prototypical case of [+oblique] [−A] NPs in Dyirbal antipassive constructions.

4.4   *Summary of clause-internal packaging constructions*

We can now summarize our discussion of passive, antipassive and applicative constructions as in Table 7.4. Passives and antipassives are defined universally as blocks on the normal linking patterns in lexical entries. They derive new lexical forms through overt derivational affixation or other morphological processes in the verb or verbal complex which prohibit the unmarked linking patterns of the arguments [+A] and [−A], so that one of these fails to link to a [−oblique] function, [+A] for a passive derivation and [−A] for an antipassive. This, de facto, forces these arguments to be realized as [+oblique], and if this is all that occurs or is intended, then a backgrounding construction results. If, however, the point of blocking the linking of either the [+A] or [−A] to [−oblique] is to meet certain clausal syntactic constraints, particularly those germane to pivots, then a foregrounding construction results, presenting the remaining [−oblique] argument as the sole one of a derived intransitive verb and thereby conferring the syntactic properties of pivots upon it. It should be pointed out that, while

Table 7.4 *Summary of voice constructions*

	Passive	Antipassive	Applicative
Core definition	$[+\text{A}] \neq [-\text{oblique}]$	$[-\text{A}] \neq [-\text{oblique}]$	$[+\text{oblique}] = [-\text{oblique}]$
Backgrounding	$[+\text{A}] = [+\text{oblique}]$	$[-\text{A}] = [+\text{oblique}]$	
Foregrounding	$[-\text{A}] = [+\text{pivot}]$	$[+\text{A}] = [+\text{pivot}]$	

foregrounding passives and antipassives are diagnostic and prototypical of pivot languages, they may not be restricted to them. It is conceivable that pivotless languages could possess foregrounding passive constructions, if the purpose of the derivation is to get the $[-\text{oblique}]$ $[-\text{A}]$ argument to assume some of the grammatical properties of the $[+\text{A}]$ argument, the most prominent argument in the argument structure (some Bantu languages may indeed be pivotless languages with exactly this property). When this occurs with a marker in the verbal expression we can say it is a passive construction. Such a derivation, of course, would not be necessary via foregrounding antipassives, because $[+\text{A}]$ arguments, by a universal algorithm, are already the most prominent argument in the argument structure and the controller and target of many constructions.

Applicatives differ from passives and antipassives in that they are not blocks on the linking of normally subcategorized $[+\text{A}]$ or $[-\text{A}]$ arguments to $[-\text{oblique}]$, but the introduction of normally $[+\text{oblique}]$ non-subcategorized arguments into the argument structure of a derived verb and linked to $[-\text{oblique}]$ status. The arguments introduced by the applicative affixes are realized as $[-\text{A}]$. This is normally straightforward in most types of languages and has little systematic effect on the lexical entries of the derived verbs, but in asymmetrical ergative–absolutive pivot languages, the one pivot per clause constraint requires the effect of a subsidiary antipassive to force the $[-\text{A}]$ argument of the underived verb into $[+\text{oblique}]$ status, so there is only one $[-\text{oblique}]$ $[-\text{A}]$ argument for the derived applicative verb, and thus one pivot.

5   On clause-external packaging options: topicalizations, left dislocations, and right dislocations

In section 3.3 we motivated a distinction between pivot, and topic and particularly the existence of constructions in which the topic of the sentence was not the pivot, such as *Soukous, I regard as the best dance music around*. The claim was that the topic in such sentences was external to the clause, not a direct constituent of it. Languages often have a number of these clause-external packaging constructions, and these will be the subject of this section.

Topicalizations and left-dislocation constructions are superficially similar: each presents a topic NP juxtaposed immediately to the left of the clause. They are distinguished by the presence in left dislocations of a pronominal element within the clause referring to the topic NP; this is absent in topicalizations. The following examples illustrate both types, (179) for topicalizations and (180) for left dislocations:

(179)   (a) That movie I wouldn't see if you gave me a free ticket
         (b) That dish, I haven't tried
         (c) For Egbert, I would do anything

(180)   (a) Turtles, they make the greatest pets
         (b) Thai cooking, I find it irresistible
         (c) Mary, I went to university with her

Note that only left dislocations are available to [+pivot] NPs; topicalizations would leave the pivot position unoccupied, which is ungrammatical in English:

(181)   *Turtles, make the greatest pets.

These constructions foreground a given NP by making it topic, but differ structurally from foregrounding passives and antipassives in that there is no distinctive mark in the verbal expression, and the topic NP is external to the clause, not internal as with the different voice constructions. This structural difference is tied to a functional one. Whereas foregrounding passives and antipassives are typically used to indicate the *continuity* of reference of a topic or pivot NP over clauses, topicalization and left dislocations have the opposite function: Lambrecht (1994) points out that the function of these constructions is to introduce a new topic or re-introduce one that was introduced previously but has not been mentioned for some clauses; in other words, these constructions point up *discontinuity* of topic. This usage is well illustrated in the mini-fable drawn from Givón (1976):

(182)   Once there was a *wizard*. *He* was very wise, rich, and was married to a beautiful witch. They had two sons. The first was tall and brooding, he spent his days in the forest hunting snails, and his mother was afraid of him. The second was short and vivacious, a bit crazy but always game. Now *the wizard*, *he* lived in Africa.

Note the final sentence contains a left-dislocated NP, *the wizard*. This is because the wizard, while already mentioned in the text, has not been a participant or topic of several previous clauses, the topics of which were his sons and their mother. Consequently, a left-dislocation construction is proper to re-introduce him as topic of the final clause.

Topicalizations and left-dislocation constructions are quite wide-spread among the languages of the world. Tagalog has pervasive and common topicalization constructions. Interestingly, it distinguishes between topicalizations involving pivots, which require a particle *ay* (183b, e), and those of [+oblique] arguments, which do not (183c); [−oblique] [−pivot] NPs cannot normally be topicalized (183d):

(183)
(a) nag-bigay   ng   isda ang   lalake sa   bata
    pivot=[+A]-give [−oblique] fish [+pivot] man   [+oblique] child
    'The man gave fish to the child'

(b) ang   lalake ay   nag-bigay   ng   isda
    [+pivot] man   TOPIC pivot=[+A]-give [−oblique] fish
    sa   bata
    [+oblique] child
    'The man, (he) gave fish to the child'

(c) sa   bata nag-bigay   ng   isda ang   lalake
    [+oblique] child pivot=[+A]-give [−oblique] fish [+pivot] man
    'the child, the man gave fish'

(d) *ng   isda nag-bigay   ang   lalake sa   bata
    [−oblique] fish pivot=[+A]-give [+pivot] man   [+oblique] child
    'Fish, the man gave to the child'

(e) ang   isda 'y   i-bi-bigay   ng   lalake
    [+pivot] fish TOPIC pivot=[−A]-FUT-give [−oblique] man
    sa   bata
    [+oblique] child
    'The fish, the men will give (it) to the child'

The restrictions against sentences like (183d) make sense in view of the fact that, prototypically, pivots are topics in Tagalog, and, as a symmetrical language, any [−oblique] NP can freely be pivot and topic. Hence, if a [−oblique] NP is to be topic, the language mandates that it also be pivot as in (183e).

Left-dislocation constructions are perhaps even more common cross-linguistically than topicalization. Jakaltek (Craig (1977)) illustrates them well:

(184)   Jakaltek
          (a) x-ɸ-s-mak   naj Pel   ix   Malin
              ASP-3SG.ABS-3SG.ERG-hit CLSFR Peter CLSFR Mary
                     [−A]   [+A]
              'Peter hit Mary'

(b) naj   Pel   x-ɸ-s-mak   naj   ix   Malin
CLSFR Peter ASP-3SG.ABS-3SG.ERG-hit CLSFR CLSFR Mary
                        [−A]   [+A]
'Peter, he hit Mary'

(c) ix   Malin x-ɸ-s-mak   naj   Pel   ix
CLSFR Mary   ASP-3SG.ABS-3SG.ERG-hit CLSFR Peter CLSFR
                        [−A]   [+A]
'Mary, Peter hit her'

In all these Jacaltec left-dislocation constructions, the external topic NP is refer-enced within the clause by a bound verbal pronominal affix and a free pronoun, which is here represented by CLSFR for 'classifier'.

It needs to be borne in mind that topic NPs are not necessarily realized in clause-external position. Besides being pivots in pivot languages, topics can in many languages be marked *in situ* within the clause with a special topic affix or particle; this is very common in Papuan languages, for example Tauya:

(185) (a) 'i   fanu-ni   fena'a-na   yau-a-'a
          DEM man-ERG woman-TOPIC see-3SG[+A]-INDIC
          'The woman, that man saw (her)'

      (b) 'i   fanu-na   pai yau-a-'a
          DEM man-TOPIC pig see-3SG[+A]-INDIC
          'That man, (he) saw the pig'   MacDonald (1994)

In addition to NPs being juxtaposed externally to the left margin of clauses in left dislocations, there are many languages which can also juxtapose them to the right margin. These are called right-dislocation or antitopic (Lambrecht (1994)) constructions. Examples in various languages include:

(186) (a) He's a good friend, Sam.
      (b) Yimas
          mum   pu-n-mampi-awkura-mpi-api-k,   paympan
          3PL   3PL[−A]-3SG[+A]-again-gather-SEQ-put in-IRR eagle
          'He again gathered them and put them inside, the eagle'

      (c) Cayuga
          kye·' sakáeyǫ',   kashehawáhkshǫ'
          then   3PL.return 3PL.POSS.daughters
          'Then they returned, your daughters'   Mithun (1992)

In the free word order languages, Yimas and Cayuga, this clause-external right-dislocation construction is distinguished from the typically freely variant order of clause-internal constituents by prosodic factors. In Yimas, for instance, the

right-dislocated NP is set off by a distinct pause from the clause itself and has a marked low falling pitch contour.

Lambrecht (1994) discusses the function of right-dislocation constructions. Unlike left dislocations, they commonly indicate already-mentioned referents that are well established in the immediate discourse, but whose role may be shifting between clauses. Or they may be used to disambiguate the referents of free or bound pronominals from a range of potential referents already established in the discourse. Unlike left dislocations, they are never used to introduce completely new topics; this is because the right-dislocated NP is mentioned in the clause through bound or free pronominals before it is ever uttered, and pronominals are characteristic of already given material. Right-dislocated NPs are also typically of low pitch prosodically; this prohibits them from signalling contrastive topics, which are marked by high falling pitch like focussed NPs. Again the contrastive function is a speciality of left dislocations. Thus, while both left and right dislocations present NPs externally to the clause, at the left and right margins respectively, they serve quite distinct functions, a fact no doubt linked to their different structural positions.

6   Suggestions for further reading

For more depth on the concepts of [+A] and [−A] and the relative accessibility of semantic role types to these perspective macroroles see Foley and Van Valin (1984), Jackendoff (1990), and Dowty (1991). Klaiman (1991) has a good discussion of the concept of agency and the related notions of control and volitionality. For various approaches to the concept of argument structure, have a look at Pinker (1989), Grimshaw (1990), Goldberg (1994), Mohanan (1994) and Manning (1996). The concept of pivot and a typology of pivot types is well covered in Dixon (1994); other valuable sources are Keenan (1976c) and Schachter (1976, 1977). The pragmatic notions of topic and focus are fully explored in Lambrecht (1994); for some typologically interesting syntactic reflexes of these notions, look at Bresnan and Mchombo (1987) and the articles in Downing and Noonan (1995). Du Bois (1987) is a very valuable study of the behaviour of these pragmatic notions in ergative languages, with important general typological implications. The literature on voice, passives and antipassives is voluminous, but excellent beginning sources are Shibatani (1988), Klaiman (1991), and Fox and Hopper (1994). Finally, for clause-external packaging options, consult Prince (1978, 1981), Davison (1984), Rochemont, and Culicover (1990) and Aissen (1992).

Bibliography

Abbott, Miriam. 1991. Macushi. In Desmond C. Derbyshire and Geoffrey K. Pullum (eds.), *Handbook of Amazonian Languages*, vol. III. Berlin: Mouton de Gruyter, 23–160.

Abdulla, Jamal Jalal, and Ernest N. McCarus. 1967. *Kurdish Basic Course*. Ann Arbor: University of Michigan Press.

Adams, Karen L., and Nancy F. Conklin. 1974. On the numerical classifier in Thai. Paper presented at the Seventh International Conference on Sino-Tibetan Language and Linguistic Studies. Atlanta, Georgia.

Aikhenvald, Alexandra Y. 2003. Imperatives and other commands. Unpublished ms. La Trobe University.

Aissen, Judith. 1992. Topic and focus in Mayan. *Language* 68, 43–80.

   1997. The syntax of obviation. *Language* 73, 705–50.

   1999. Agent focus and inverse in Tzotzil. *Language* 75, 451–85.

Akmajian, Adrian. 1984. Sentence types and the form-function fit. *Natural Language and Linguistic Theory* 2, 1–24.

Alexander, Ruth Mary. 1988. A syntactic sketch of Ocotepec Mixtec. In C. Henry Bradley and Barbara E. Hollenbach (eds.), *Studies in the Syntax of Mixtecan Languages,* vol. I. Dallas: Summer Institute of Linguistics and The University of Texas at Arlington, 151–304.

Alsina, Alex. 1993. Predicate composition: a theory of syntactic function alternations. Unpublished Ph.D. dissertation. Stanford University.

   1996a. Passive types and the theory of argument asymmetries. *Natural Language and Linguistic Theory* 14, 673–723.

   1996b. *The Role of Argument Structure in Grammar: Evidence from Romance*. Stanford: CSLI.

Alsina, Alex, Joan Bresnan, and Peter Sells (eds.). 1997. *Complex Predicates*. Stanford: CSLI.

Altmann, H. 1984. Zur Problematik der Konstitution von Satzmodi als Formtypen. In Jörg Meibauer (ed.), *Satzmodus zwischen Grammatik und Pragmatik*. Tübingen: Niemeyer, 22–56.

   1993. Satzmodi. In Joachim Jacobs, Arnim von Stechow, Wolfgang Sternefeld, and Theo Venneman (eds.), *Syntax. Ein internationales Handbuch zeitgenössischer Forschung*. Berlin: de Gruyter, 1006–29.

Andersen, Torben. 1988. Ergativity in Päri, a Nilotic OVS language. *Lingua* 75, 289–324.

Anderson, Stephen R. 1976. On the notion of subject in ergative languages. In Charles N. Li (ed.), *Subject and Topic*. New York: Academic Press, 1–23.

Andrews, Avery D. 1985. The major functions of the noun phrase. In Timothy Shopen (ed.), *Language Typology and Syntactic Description*, vol. I: *Clause Structure*. Cambridge: Cambridge University Press, 62–154.
  1990. Case-structures and control in Modern Icelandic. In Joan Maling and Annie Zaenen (eds.), *Modern Icelandic Syntax*. New York: Academic Press, 187–234.
  2001. Noncanonical A/S-marking in Icelandic. In Alexandra Aikhenvald and R. M. W. Dixon (eds.), *Noncanonical Case Marking of Subjects and Objects*. Amsterdam: Benjamins, 85–109.
Aristar, Anthony Rodrigues. 1991. On diachronic sources and synchronic pattern: an investigation into the origin of linguistic universals. *Language* 67, 1–33.
Asher, R. E., and T. C. Kumari. 1997. *Malayalam*. London: Routledge.
Asyik, Abdul Gani. 1987. A contextual grammar of acehnese sentences. Ph.D. thesis. University of Michigan.
Austin, John. 1962. *How to do Things with Words*. Oxford: Clarendon.
Austin, Peter K. 1981a. *A Grammar of Diyari*. Cambridge: Cambridge University Press.
  1981b. Switch-reference in Australia. *Language* 57, 309–33.
Austin, Peter K., and Joan W. Bresnan. 1997. Causatives and applicatives in Australian Aboriginal languages. In Kazuto Matsumura and Tooru Hayashi (eds.), *The Dative and Related Phenomena*. Tokyo: Hitsuji Shobo, 165–225.
Bach, Emmon, and Robert T. Harms (eds.). 1968. *Universals in Linguistic Theory*. New York: Holt, Rinehart, and Winston.
Baker, Carl Lee. 1970. Notes on the description of English questions. *Foundations of Language* 6, 197–219.
Baker, Mark C. 1988. *Incorporation*. Chicago: University of Chicago Press.
Bakhtin, M. 1981. *The Dialogic Imagination: Four Essays*. Austin: University of Texas Press.
Barnes, J. 1984. Evidentials in the Tuyuca verb. *International Journal of American Linguistics* 50, 255–71.
Bascom, Burton. 1982. Northern Tepehuan. In Ronald W. Langacker (ed.), *Uto-Aztecan Grammatical Sketches*. Studies in Uto-Aztecan Grammar 3. Dallas: Summer Institute of Linguistics, 267–393.
Bateson, M. C. 1967. *Arabic Language Handbook*. Washington: Center for Applied Linguistics.
Bauer, Winifred. 1993. *Maori*. Descriptive Grammars. London: Routledge.
Bell, Sarah J. 1976. Cebuano subjects in two frameworks. Ph.D. thesis. Massachussetts Institute of Technology. (Distributed by the Indiana University Linguistics Club.)
Besnier, Nico. 2000. *Tuvaluan. A Polynesian Language of the Central Pacific*. London: Routledge.
Bhat, D. N. S. 1991. *Grammatical Relations: The Evidence Against Their Necessity and Universality*. London: Routledge.
  2000. Word classes and sentential functions. In P. M. Vogel and B. Comrie (eds.), *Approaches to the Typology of Word Classes*. Berlin: Mouton, 47–63.
Bhatia, Tej K. 1993. *Punjabi*. Descriptive Grammars. London: Routledge.
Binh, D. T. 1971. *A Tagmemic Comparison of the Structure of English and Vietnamese Sentences*. The Hague: Mouton.
Birk, D. B. W. 1976. *The Malakmalak Language, Daly River (Western Arnhem Land)*. Pacific Linguistics Series B 45. Canberra: The Australian National University.

Bittner, Maria. 1994. *Case, Scope and Binding*. Dordrecht: Academic.

Black, K., and Mrs. K. Black. 1971. *The Moro Language Grammar and Dictionary*. Linguistic Monograph Series 6. Khartoum: Sudan Research Unit, Faculty of Arts.

Blake, Barry J. 1990. *Relational Grammar*. London: Routledge.

  1994. *Case*. Cambridge: Cambridge University Press.

Bloomfield, Leonard. 1917. *Tagalog Texts with Grammatical Analysis*. University of Illinois Studies in Language and Literature 3. Urbana: University of Illinois.

  1934. *Plains Cree Texts*. American Ethnological Society Publications 16. New York: American Ethnological Society.

  1958. *Eastern Ojibwa: Grammatical Sketch, Texts, and Word List*. Ann Arbor: University of Michigan Press.

Boas, Franz. 1917. *Grammatical Notes on the Language of the Tlingit Indians*. Anthropological Publications 8, 1. Philadelphia: University of Pennsylvania Museum.

Bolinger, Dwight L. 1967. Adjectives in English: attribution and predication. *Lingua* 18, 1–34.

Borgman, Donald M. 1990. Sanuma. In Desmond C. Derbyshire and Geoffrey K. Pullum (eds.), *Handbook of Amazonian Languages*, vol. II. Berlin: Mouton de Gruyter, 15–248.

Bowden, John. 1997. Taba (Makian Dalam): description of an Austronesian language of Eastern Indonesia. Ph.D. thesis. University of Melbourne.

Bresnan, Joan W. (ed.). 1982. *The Mental Representation of Grammatical Relations*. Cambridge, MA: MIT Press.

  1994. Locative inversion and the architecture of universal grammar. *Language* 70, 72–131.

  2001. *Lexical-Functional Syntax*. Oxford: Blackwell.

Bresnan, Joan W., and Sam Mchombo. 1987. Topic, pronoun and agreement in Chichêwa. *Language* 63, 741–82.

Bresnan, Joan W., and Lioba Moshi. 1990. Object asymmetries in comparative Bantu syntax. *Linguistic Inquiry* 21, 147–185. Reprinted in Sam A. Mchombo (ed.), *Theoretical Aspects of Bantu Grammar*. Stanford: CSLI, 47–91.

Brown, Lea. 2001. A grammar of Nias Selatan. Ph.D. dissertation. University of Sydney.

Burzio, Luigi. 1986. *Italian Syntax: A Government–Binding Approach*. Dordrecht: Reidel.

Carlson, Robert. 1994. *A Grammar of Supyire*. Berlin: Mouton de Gruyter.

Catford, J. C. 1976. Ergativity in Caucasian languages. Proceedings of the Northeastern Linguistics Society. *Recherches Linguistiques à Montréal* 6, 37–48.

Chafe, Wallace L. 1976. Givenness, contrastiveness, definiteness, subjects, topics and point of view. In Charles N. Li (ed.), *Subject and Topic*. New York: Academic Press, 25–55.

  1990. Use of defocusing pronominal prefixes in Caddo. *Anthropological Linguistics* 32, 57–68.

Chafe, Wallace L., and Johanna Nichols (eds.). 1986. *Evidentiality: The Linguistic Coding of Epistemology*. Norwood, NJ: Ablex Publishing Corporation.

Chang, Suk-Jin. 1996. *Korean*. Amsterdam: Benjamins.

Chao, Y.-R. 1968. *A Grammar of Spoken Chinese*. Berkeley: University of California Press.

Cheng, Lisa Lai-Shen. 1997. *On the Typology of Wh-Questions*. New York: Garland.

Chhangte, Lalnunthangi. 1989. The grammar of simple clauses in Mizo. In David Bradley (ed.), *South-East Asian Syntax*. Pacific Linguistics, Series A 77 (Papers in South-East Asian Linguistics 11). Canberra: The Australian National University, 93–174.

Chisholm, William S., Louis T. Milic, and John A. C. Greppin (eds.). 1984. *Interrogativity: A Colloquium on the Grammar, Typology and Pragmatics of Questions in Seven Diverse Languages*. Amsterdam: Benjamins.

Choi, Hye-Won. 1999. *Optimizing Structure in Context: Scrambling and Information Structure*. Stanford: CSLI.

Chomsky, Noam. 1957. *Syntactic Structures*. The Hague: Mouton.

Christaller, Revd J. G. 1875. *A Grammar of the Asante and Fante Language called Tshi (Chwee, Twi)*. Gold Coast: Basel German Evangelical Mission. (Reprinted 1964, Ridgewood, NJ: Gregg Press.)

Chung, Sandra. 1976. An object-creating rule in Bahasa Indonesia. *Linguistic Inquiry* 7, 41–87.

Clark, Herbert H. 1996. *Using Language*. Cambridge: Cambridge University Press.

Cole, Peter, and Jerrold M. Sadock (eds.). 1977. *Syntax and Semantics,* vol. VIII*: Grammatical Relations*. New York: Academic Press.

Collins, G. 1970. Two views of Kalagan grammar. Unpublished Ph.D. dissertation. Indiana University.

Comrie, Bernard. 1977. In defense of spontaneous demotion: the impersonal passive. In Peter Cole and Jerrold M. Sadock (eds.), *Syntax and Semantics,* vol. VIII*: Grammatical Relations*. New York: Academic Press, 47–58.

  1978. Ergativity. In Winfred P. Lehmann (ed.), *Syntactic Typology*. Austin: University of Texas Press, 329–394.

  1979. Russian. In T. Shopen (ed.), *Languages and their Speakers*. Cambridge, MA: Winthrop, 91–151.

  1981. Ergativity in Kalaw Lagaw Ya. *Australian Journal of Linguistics* 11, 1–42.

  1987. Russian. In Timothy Shopen (ed.), *Languages and their Status*. Philadelphia: University of Pennsylvania Press, 91–151.

  2005. Alignment of case marking. In M. Haspelmath *et al.* (eds.), *The World Atlas of Language Structures*. Oxford: Oxford University Press, 398–405.

Comrie, Bernard, and Maria Polinsky. In press. On the grammar of evidentiality. Witnessed and unwitnessed verb forms in Tsez. In Zlata Guentcheva (ed.), *Studies in Evidential Constructions*. Amsterdam: Benjamins.

Conrad, Robert J., and Kepas Wogiga. 1991. *An Outline of Bukiyip Grammar*. Pacific Linguistics, Series C 113. Canberra: The Australian National University.

Cooreman, Ann. 1987. *Transitivity and Discourse Continuity in Chamorro Narratives*. Berlin: Mouton de Gruyter.

  1988. The antipassive in Chamorro: variations on the theme of transitivity. In Masayoshi Shibatani (ed.), *Passive and Voice*. Amsterdam: Benjamins, 561–93.

  1994. A functional typology of antipassives. In Barbara Fox and Paul J. Hopper (eds.), *Voice: Form and Function*. Amsterdam: John Benjamins, 49–87.

Cooreman, Ann, Barbara Fox, and T. Givón. 1984. The discourse definition of ergativity. *Studies in Language* 8, 1–34.

Corston, S. H. 1996. *Ergativity in Roviana, Solomon Islands*. Pacific Linguistics, Series B 113. Canberra: The Australian National University.

Courtenay, K. 1976. Ideophones defined as a phonological class: the case of Yoruba. *Studies in African Linguistics* Supplement 6, 13–26.

Craig, Colette. 1977. *The Structure of Jacaltec*. Austin: University of Texas Press.

Crazzolara, P. (1955) *A Study of the Acooli Language*. Oxford: Oxford University Press.

Croft, William. 1984. Semantic and pragmatic correlates to syntactic categories. In David Testen, Veena Mishra, and Joseph Drago (eds.), *Papers from the Parasession on Lexical Semantics*. Chicago: Chicago Linguistics Society, 53–70.

1991. *Typology and Universals*. Cambridge: Cambridge University Press.

1994. Speech act classification, language typology and cognition. In Svas L. Tsohatzidis (ed.), *Foundations of Speech Acts Theory. Philosophical and Linguistics Perspectives*. London: Routledge, 460–77.

Crowley, Terry 1978. *The Middle Clarence Dialects of Bandjalang*. Canberra: Australian Institute of Aboriginal Studies.

1981. The Mpakwithi dialect of Anguthimri. In R. M. W. Dixon and Barry J. Blake (eds.), *Handbook of Australian Languages*, vol. II. Amsterdam: John Benjamins, 146–94.

Curme, G. O. 1935. Parts of speech and accidence. In G. O. Curme and H. Kurath, *A Grammar of the English Language*, vol. II: *Parts of Speech and Accidence*. Boston: D. C. Heath and Co.

Curnow, Timothy J. 1997. A grammar of Awa Pit (Cuaiquer): an indigenous language of South-western Colombia. Ph.D. thesis. Australian National University.

Dahlstrom, Amy. 1986. Plains Cree morphosyntax. Unpublished Ph.D. dissertation. University of California, Berkeley.

1991. *Plains Cree Morphosyntax*. New York: Garland Publishing.

Dalrymple, Mary, Ronald M. Kaplan, John T. Maxwell, and Annie Zaenen. 1995. *Formal Issues in Lexical-Functional Grammar*. Stanford: CSLI.

Davies, Eirlys. 1986. *The English Imperative*. London: Croom Helm.

Davies, John. 1981. *Kobon*. Lingua Descriptive Studies 3. Amsterdam: North-Holland.

Davies, William D. 1986. *Choctaw Verb Agreement and Universal Grammar*. Dordrecht: D. Reidel

Davison, A. 1984. Syntactic markedness and the definition of sentence topic. *Language* 60, 797–846.

Dayley, J. 1985. Voice in Tzutujil. In Johanna Nichols and Anthony C. Woodbury (eds.), *Grammar Inside and Outside the Clause*. Cambridge: Cambridge University Press, 192–226.

DeLancey, Scott. 1981. An interpretation of split ergativity and related patterns. *Language* 57, 626–57.

1997. Mirativity: the grammatical marking of unexpected information. *Linguistic Typology* 1, 33–52.

De Vries, Lourens. 1993. *Forms and Functions in Kombai, an Awyu Language of Irian Jaya*. Pacific Linguistics, Series B 108. Canberra: The Australian National University.

Dench, Alan Charles. 1995. *Martuthunira: A Language of the Pilbara Region of Western Australia*. Pacific Linguistics, Series C 125. Canberra: The Australian National University.

Derbyshire, Desmond. 1979. *Hixkaryana*. Lingua Descriptive Studies 1. Amsterdam: North-Holland.

Dez, Jacques. 1980. *La syntaxe du Malgache*. Lille, France: Atelier de reproduction des thèses.

Diessel, Holger. 1999. *Demonstratives: Form, Function and Grammaticalization*. Amsterdam: John Benjamins.

Dimmendaal, G. J. 1983. *The Turkana Language*. Dordrecht: Foris Publications Holland.

Dixon, R. M. W. 1972. *The Dyirbal Language of North Queensland*. Cambridge: Cambridge University Press.

    1977a. *A Grammar of Yidiny*. Cambridge: Cambridge University Press.

    1977b. Where have all the adjectives gone? *Studies in Language* 1, 19–80. Reprinted in Dixon, *Where Have All the Adjectives Gone? and Other Essays in Semantics and Syntax*. New York: Mouton, 1982.

    1977c. The syntactic development of Australian languages. In Charles N. Li (ed.), *Mechanisms of Syntactic Change*. New York: Academic Press, 365–415.

    1979. Ergativity. *Language* 55, 59–138.

    1980. *The Languages of Australia*. Cambridge: Cambridge University Press.

    1981. Grammatical reanalysis: an example of linguistic change from Warrgamay (North Queensland). *Australian Journal of Linguistics* 1, 91–112.

    1988. *A Grammar of Boumaa Fijian*. Chicago: University of Chicago Press.

    1991. *A New Approach to English Grammar, on Semantic Principles*. Oxford: Clarendon Press.

    1994. *Ergativity*. Cambridge: Cambridge University Press.

    1995. Complement clauses and complement strategies. In Frank R. Palmer (ed.), *Grammar and Meaning: Essays in Honour of Sir John Lyons*. Cambridge: Cambridge University Press, 175–220.

Doke, C. M. 1935. *Bantu Linguistic Terminology*. London: Longmans, Green.

Donaldson, Tamsin. 1980. *Ngiyambaa: The Language of the Wangaaybuwan*. Cambridge: Cambridge University Press.

Donohue, Mark. 1999. *A Grammar of Tukang Besi*. Berlin: Mouton de Gruyter.

Dorais, Louis J. 1978. *Lexique analytique du vocabulaire Inuit moderne au Québec-Labrador*. Quebec: Presses de l'Université Laval.

Downing, P., and M. Noonan (eds.). 1995. *Word Order in Discourse*. Amsterdam: Benjamins.

Dowty, D. 1991. Thematic proto-roles and argument selection. *Language* 67, 547–619.

Driem, George van. 1987. *A Grammar of Dumi*. Mouton Grammar Library 10. Berlin: Mouton de Gruyter.

Dryer, Matthew S. 1983. Indirect objects in Kinyarwanda revisited. In David M. Perlmutter (ed.), *Studies in Relational Grammar*, vol. I. Chicago: University of Chicago Press, 129–140.

    1985. Tlingit: an object-initial language? *Canadian Journal of Linguistics* 30, 1–13.

    1986. Primary objects, secondary objects and antidative. *Language* 62, 808–845.

    1988. Object–verb order and adjective–noun order: dispelling a myth. *Lingua* 74, 185–217.

    1992. The Greenbergian word order correlations. *Language* 68, 81–138.

    1994. The discourse function of the Kutenai inverse. In T. Givón (ed.), *Voice and Inversion*. Amsterdam: John Benjamins, 65–99.

    1997a. On the six-way word order typology. *Studies in Language* 21, 69–103.

1997b. Passive vs. indefinite subject construction in Cree. In David Pentland (ed.), *Papers of the Twenty-Sixth Algonquian Conference*. Winnipeg: University of Manitoba.

2005. Order of subject, object, and verb. In M. Haspelmath, *et al.* (eds.), *The World Atlas of Language Structures*. Oxford: Oxford University Press, 330–3.

Du Bois, John. 1987. The discourse basis of ergativity. *Language* 63, 805–55.

Du Feu, Veronica. 1996. *Rapanui*. Descriptive Grammars. London: Routledge.

Duranti, A., and E. Byarushengo. 1977. On the notion of 'direct object' in Haya. In Ernest R. Byarushengo *et al.* (eds.), *Haya Grammatical Structures: Phonology, Grammar, Discourse.* Southern California Occasional Papers in Linguistics 6. Los Angeles: Dept. of Linguistics, University of Southern California, 45–71.

Durie, Mark. 1985. *A Grammar of Acehnese, on the Basis of a Dialect of North Acehnese*. Dordrecht: Foris.

1987. Grammatical relations in Acehnese. *Studies in Language* 11, 365–99.

1988. Preferred argument structure in an active language. *Lingua* 74, 1–25.

Dziwirek, Katarzyna, Patrick Farrell and Errapel Mejías Bikandi (eds.). 1990. *Grammatical Relations: A Cross-Theoretical Perspective*. Stanford: CSLI.

Ebermann, Erwin. 1986. Die Sprache der Mauka. Unpublished dissertation: VWGO Vienna.

Emanatian, M. 1988. The Nootka passive revisited. In W. Shipley (ed.), *In Honor of Mary Haas: From the Haas Festival Conference on Native American Linguistics*, Berlin: Mouton de Gruyter, 265–291.

England, Nora. 1983a. *A Grammar of Mam, a Mayan Language*. Austin: University of Texas Press.

1983b. Ergativity in Mamean (Mayan) languages. *International Journal of American Linguistics* 49, 1–19.

1988. Mam voice. In Masayoshi Shibatani (ed.), *Passive and Voice*. Amsterdam: Benjamins, 525–45.

Evans, Nicholas D. 1995. *A Grammar of Kayardild*. Berlin: Mouton de Gruyter.

Ezard, Brian. 1997. *A Grammar of Tawala: An Austronesian Language of the Milne Bay Area, Papua New Guinea*. Pacific Linguistics, Series C 137. Canberra: The Australian National University.

Facundes, Sidney Da S. 2000. The language of the Apurinã people of Brazil (Maipure/Arawakan). Ph.D. dissertation. State University of New York at Buffalo.

Faltz, Leonard M. 1977. Reflexivization: a study in universal syntax. Unpublished Ph.D. dissertation. University of California, Berkeley.

Feldman, Harry. 1986. *A Grammar of Autuw*. Pacific Linguistics, Series B 94. Canberra: The Australian National University.

Fillmore, Charles J. 1965. *Indirect Object Constructions in English and the Ordering of Transformations*. The Hague: Mouton.

1968. The case for case. In Emmon Bach and Robert T. Harms (eds.), *Universals in Linguistic Theory*. New York: Holt, Rinehart, and Winston, 1–88.

1977. The case for case reopened. In Peter Cole and Jerrold M. Sadock (eds.), *Syntax and Semantics*, vol. VIII: *Grammatical Relations*. New York: Academic Press, 59–82.

Foley, William A. 1986. *The Papuan Languages of New Guinea*. Stanford: Stanford University Press.

Foley, William A., and Michael Olson. 1985. Clausehood and verb serialization. In Johanna Nichols and Anthony C. Woodbury (eds.), *Grammar Inside and Outside the Clause*. Cambridge: Cambridge University Press, 17–60.

Foley, William A., and R. Van Valin. 1984. *Functional Syntax and Universal Grammar*. Cambridge: Cambridge University Press.

Fortescue, Michael. 1984. *West Greenlandic*. London: Croom Helm.

Fox, Barbara, and Paul J. Hopper (eds.). 1994. *Voice: Form and Function*. Amsterdam: John Benjamins.

Frachtenberg, Leo J. 1922. Coos. In Franz Boas (ed.), *Handbook of American Indian Languages, Part II*. Smithsonian Institution Bureau of American Ethnology Bulletin 40. Washington: Government Printing Office, 297–429.

Franklin, Karl James. 1971. *A Grammar of Kewa, New Guinea*. Pacific Linguistics, Series C 16. Canberra: The Australian National University.

Frazier, Lynn. 1979. Parsing and constraints on word order. *University of Massachusetts Occasional Papers in Linguistics* 5, 177–98.

Fries, Charles. 1952. *The Structure of English*. New York: Harcourt Brace.

Gary, Judith O., and Edward L. Keenan. 1977. On collapsing grammatical relations in Universal Grammar. In Peter Cole and Jerrold M. Sadock (eds.), *Syntax and Semantics*, vol. VIII: *Grammatical Relations*. New York: Academic Press, 83–120.

Gawron, Jean Mark. 1986. Situations and prepositions. *Linguistics and Philosophy* 9, 327–82.

Geluykens, R. 1988. On the myth of rising intonation in polar questions. *Journal of Pragmatics* 12, 467–85.

Genetti, Carol. 1994. *A Descriptive and Historical Account of the Dolakha Newari Dialect*. Tokyo: Institute for the Study of Languages and Cultures of Asia and Africa.

Gerdts, Donna B. 1988. Antipassives and causatives in Ilokano. In Richard McGinn (ed.), *Studies in Austronesian Linguistics*. Athens, OH: Ohio University Center for International Studies, 295–321.

Gildea, Spike. 1994. Semantic and pragmatic inverse: 'inverse alignment' and 'inverse voice' in Carib of Surinam. In T. Givón (ed.), *Voice and Inversion*. Amsterdam: John Benjamins, 149–88.

Givón, T. 1972. Review of *Some Problems in Transitivity in Swahili*. *African Studies* 31, 273–7.

  1975. Serial verbs and syntactic change: Niger-Congo. In Charles N. Li (ed.), *Word Order and Word Order Change*. Austin: University of Texas Press, 47–112.

  1976. Topic, pronoun and grammatical agreement. In Charles N. Li (ed.), *Subject and Topic*. New York: Academic Press, 149–88.

  1979. *On Understanding Grammar*. New York: Academic Press.

  1984a. *Syntax: A Functional-typological Introduction,* Vol. I. Amsterdam: Benjamins.

  1984b. Ute. In William S. Chisholm *et al.* (eds.), *Interrogativity: A Colloquium on the Grammar, Typology and Pragmatics of Questions in Seven Diverse Languages*. Amsterdam: Benjamins, 215–43.

  (ed.). 1994. *Voice and Inversion*. Amsterdam: John Benjamins.

  2001. *Syntax: An Introduction*. Amsterdam: John Benjamins.

Glinert, Lewis. 1989. *The Grammar of Modern Hebrew*. Cambridge: Cambridge University Press.

Goldberg, Adele. 1994. *A Construction Grammar Approach to Argument Structure*. Chicago: University of Chicago Press.

Grace, George. 1987. *The Linguistic Construction of Reality*. London: Croom Helm.

Green, M. M., and G. E. Igwe. 1963. *A Descriptive Grammar of Igbo*. London: Oxford University Press.

Greenberg, Joseph H. 1963. Some universals of grammar with particular reference to the order of meaningful elements. In Joseph Greenberg (ed.), *Universals of Language*. Cambridge, MA: MIT Press, 73–113.

1966. Some universals of grammar with particular reference to the order of meaningful elements. In Joseph H. Greenberg (ed.), *Universals of Language*, 2nd edn. Cambridge, MA: MIT Press, 73–113.

Grimes, Charles E. 1991. The Buru language of Eastern Indonesia. Ph.D. thesis. Australian National University.

Grimes, Joseph. 1985. Topic inflection in Mapudungan verbs. *International Journal of American Linguistics* 51, 141–63.

Grimshaw, Jane. 1990. *Argument Structure*. Cambridge, MA: MIT Press.

Gruber, Jeffrey S. 1965. Studies in lexical relations. Ph.D. dissertation. MIT.

1976. *Lexical Structure in Syntax and Semantics*. Amsterdam: North-Holland. Reprint of 1965 Ph.D. dissertation.

Gudai, Darmansyah. 1988. A grammar of Maanyan, a language of Central Kalimantan. Ph.D. thesis. Australian National University.

Guilfoyle, E., H. Hung, and L. Travis. 1992. spec of IP and spec of VP: two subjects in Austronesian languages. *Natural Language and Linguistic Theory* 10, 375–414.

Gumperz, J. 1982. *Discourse Strategies*. Cambridge: Cambridge University Press.

Gunlogson, Christine. 2001. True to form: rising and falling declaratives as questions in English. Ph.D. dissertation. University of California at Santa Cruz.

Hagman, Roy S. 1977. *Nama Hottentot Grammar*. Bloomington: Indiana University Press.

Haiman, John. 1978. Conditionals are topics. *Language* 54, 564–89.

1985. *Natural Syntax: Iconicity and Erosion*. Cambridge: Cambridge University Press.

Hale, Kenneth L. 1973. Person marking in Warlpiri. In Stephen R. Anderson and Paul Kiparsky (eds.), *A Festschrift in Honor of Morris Halle*. New York: Holt, Rinehart, and Winston, 308–44.

1976. The adjoined relative clause in Australia. In R. M. W. Dixon (ed.), *Grammatical Categories in Australian Languages*. Canberra: Australian Institute of Aboriginal Studies, 78–105.

1982. Some essential features of Warlpiri main clauses. In Stephen Swartz (ed.), *Papers in Warlpiri Grammar: In Memory of Lother Jagst. Work-papers of SIL-AAB*, Series A 6. Berrimah, Australia: Summer Institute of Linguistics, 217–315.

Hall, Allen H. 1972. A study of the Thaayorre language of the Edward River Tribe, Cape York Peninsula, Queensland: being a description of the grammar. Ph.D. thesis. University of Queensland.

Hamblin, Charles L. 1987. *Imperatives*. New York: Blackwell.

Hardman, Martha J. 1986. Data-source marking in the Jaqi languages. In Wallace Chafe and Johanna Nichols (eds.), *Evidentiality: The Linguistic Coding of Epistemology*. Norwood, NJ: Ablex Publishing Corporation, 113–36.

Harper, K. 1974. Some aspects of the grammar of the Eskimo dialects of Cumberland Peninsula and North Baffin Island. *National Museum of Canada, Ethnology Division Paper*, 15.

Harriehausen, Bettina. 1990. *Hmong Njua*. Tübingen: Max Niemeyer Verlag.

Haspelmath, Martin. 1993. *A Grammar of Lezgian*. Berlin: Mouton de Gruyter.

 2005. Ditransitive constructions: the verb 'give'. In M. Haspelmath *et al*. (eds.), *The World Atlas of Language Structures*. Oxford: Oxford University Press, 426–9.

Haspelmath, Martin, Matthew Dryer, David Gil and Bernard Comrie (eds.). 2005. *The World Atlas of Language Structures*. Oxford: Oxford University Press.

Hausser, Roland. 1983. The syntax and semantics of English mood. In Ferenc Kiefer (ed.), *Questions and Answers*. Dordrecht: Reidel.

Hawkins, John A. 1983. *Word Order Universals*. New York: Academic Press.

 1984. Modifier–head or function–argument relations in phrase structure? *Lingua* 63, 107–38.

 1990. A parsing theory of word order universals. *Linguistic Inquiry* 21, 223–61.

 1994. *A Performance Theory of Order and Constituency*. Cambridge: Cambridge University Press.

Hawkinson, A., and L. Hyman. 1974. Hierarchies of natural topic in Shona. *Studies in African Linguistics* 5, 147–70.

Hayward, Dick. 1984. *The Arbore Language: A First Investigation (Including a Vocabulary)*. Hamburg: Helmut Buske Verlag.

Heath, Jeffrey. 1999. *A Grammar of Koyraboro (Koroboro) Senni*. Cologne: Rüdiger Köppe Verlag.

Hengeveld, Kees. 1992. *Non-verbal Predication: Theory, Typology, Diachrony*. Berlin: Mouton de Gruyter.

Hess, H. Harwood. 1968. *The Syntactic Structure of Mezquital Otomi*. The Hague: Mouton.

Hinds, John. 1984. Japanese. In William S. Chisholm *et al*. (eds.), *Interrogativity: A Colloquium on the Grammar, Typology and Pragmatics of Questions in Seven Diverse Languages*. Amsterdam: Benjamins, 145–88.

Hoffman, Carl. 1963. *A Grammar of the Margi Language*. London: Oxford University Press.

Holisky, Dee Ann. 1987. The case of the intransitive subject in Tsova-Tush (Batsbi). *Lingua* 71, 103–32.

Hopper, Paul J., and Sandra A. Thompson. 1984. The discourse basis for lexical categories in universal grammar. *Language* 60, 703–52.

Hoskison, J. T. 1983. A grammar and dictionary of the Gude language. Ph.D dissertation. The Ohio State University.

Howard, I., and A. Niyekawa-Howard. 1976. Passivization. In Masayoshi Shibatani (ed.), *Japanese Generative Grammar*. Syntax and Semantics 5. New York: Academic Press, 201–37.

Huddleston, Rodney. 1994. The contrast between interrogatives and questions. *Journal of Linguistics* 30, 411–39.

Hutchison, John P. 1976. Aspects of Kanuri syntax. Ph.D. dissertation. Indiana University. Ann Arbor: University Microfilms.

Hyman, Larry, and Alessandro Duranti. 1982. On the object relation in Bantu. In P. Hopper and S. Thompson (eds.), *Studies in Transitivity*. Syntax and Semantics 15. New York: Academic Press, 217–39.

Inman, M. V. 1993. Semantics and pragmatics of colloquial singhala involitive verbs. Ph.D. dissertation. Stanford University.

Jackendoff, Ray S. 1972. *Semantic Interpretation in Generative Grammar*. Cambridge, MA: MIT Press.

1990. *Semantic Structures*. Cambridge, MA: MIT Press.

Jacobsen, W. H., Jr. 1976. Noun and verb in Nootkan. Paper presented at the Victoria Conference on Northwestern Languages. Published 1979 in Barbara S. Efrat (ed.), *The Victoria Conference on Northwestern Languages*. Victoria: British Columbia Provincial Museum, 83–153.

Jaggar, P. 1981. Some unusual lexical passives in Hausa. Unpublished MA thesis. University of California at Los Angeles.

Jauncey, Dorothy. 1997. A grammar of Tamambo, the language of Western Malo. Ph.D. thesis. The Australian National University.

Jelinek, Eloise, and Richard Demers. 1983. The agent hierarchy and voice in some Coast Salish languages. *International Journal of American Linguistics* 49, 167–85.

1994. Predicates and pronominal arguments in Straits Salish. *Language* 70, 697–736.

Jespersen, O. 1924. *The Philosophy of Grammar*. London: Allen and Unwin.

Johanson, Lars, and Bo Utas (eds.). 2000. *Evidentials*. Berlin: Walter de Gruyter.

Jones, Wendell, and Paula Jones. 1991. *Barasano Syntax*. Dallas: Summer Institute of Linguistics and The University of Texas at Arlington.

Jónsson, Jóhannes Gísli. 1996. *Clausal Architecture and Case in Icelandic*. Amherst, MA: University of Massachussetts.

Joshi, Smita. 1989. Logical subject in Marathi grammar and the predicate argument structure. In E. Fee and K. Hunt (eds.), *Proceedings of the Eighth West Coast Conference on Formal Linguistics*. Stanford: CSLI, 207–19.

Kachru, Y. B., B. Kachru and K. Bhatia. 1976. The notion 'subject', a note on Hindi–Urdu. In K. Verma Manindra (ed.), *The Notion of Subject in South Asian Languages*. Madison: University of Wisconsin, 79–108.

Kakumasu, James. 1986. Urubu-Kaapor. In Desmond C. Derbyshire and Geoffrey K. Pullum (eds.), *Handbook of Amazonian Languages*, vol. I Berlin: Mouton de Gruyter, 326–403.

Kay, Paul, and Charles J. Fillmore. 1999. Grammatical constructions and linguistic generalizations: the *What's X doing Y?* construction. *Language* 75, 1–34.

Keen, Sandra. 1983. Yukulta. In R. M. W. Dixon and Barry J. Blake (eds.), *Handbook of Australian Languages*, vol. III. Amsterdam: John Benjamins, 191–304.

Keenan, Edward L. 1972. On semantically based grammar. *Linguistic Inquiry* 3, 413–61.

1976a. Remarkable subjects in Malagasy. In Charles N. Li (ed.), *Subject and Topic*. New York: Academic Press, 247–301.

1976b. Some universals of passive in relational grammar. In *Papers from the 11th Meeting of the Chicago Linguistic Society*. Chicago: Chicago Linguistic Society, 340–52.

1976c. Towards a universal definition of subject. In Charles N. Li (ed.), *Subject and Topic*. New York: Academic Press, 303–33.

1978. The syntax of subject-final languages. In Winfred P. Lehmann (ed.), *Syntactic Typology*. Austin: University of Texas Press, 267–327.

1979. On surface form and logical form. In Braj B. Kachru (ed.), *Linguistics in the Seventies: Directions and Prospects*. Studies in the Linguistic Sciences 8. Urbana, IL: Department of Linguistics, University of Illinois. Reprinted 1985 in Edward L. Keenan, *Universal Grammar: 15 Essays*. London: Croom Helm, 375–428.

Keenan, Edward L., and Bernard Comrie. 1977. Noun phrase accessibility and universal grammar. *Linguistic Inquiry* 8, 1, 63–99.

Kenesei, István, Robert M. Vago, and Anna Fenyvesi. 1998. *Hungarian*. Descriptive Grammars. London: Routledge.

Kimenyi, Alexandre. 1980. *A Relational Grammar of Kinyarwanda*. University of California Publications in Linguistics 91. Berkeley: University of California Press.

King, Tracy Holloway. 1995. *Configuring Topic and Focus in Russian*. Stanford: CSLI.

Kirsner, R. 1976. On the subjectless pseudo-passive in standard Dutch and the semantics of background agents. In Charles N. Li (ed.), *Subject and Topic*. New York: Academic Press, 385–415.

Kiss, Katalina. 1987. *Configurationality in Hungarian*. Dordrecht: Reidel.

(ed.). 1995. *Discourse Configurational Languages*. Oxford: Oxford University Press.

Kisseberth, C., and M. Abasheikh. 1977. The object relationship in Chi-Mwi:ni, a Bantu language. In Peter Cole and Jerrold M. Sadock (eds.), *Syntax and Semantics*, vol. VIII: Grammatical Relations. New York: Academic Press, 179–218.

Klaiman, Mimi H. 1986. Semantic parameters and the South Asian linguistics area. In B. Krishnamurti, C. P. Masica, and A. K. Sinha (eds.), *South Asian Languages: Structure, Convergence and Diglossia*. Delhi: Motilal Banarsidass, 179–94

1991. *Grammatical Voice*. Cambridge: Cambridge University Press.

Klamer, Marian. 1998. *A Grammar of Kambera*. Berlin: Mouton de Gruyter.

Kleinschmidt, S. P. 1851. *Grammatik der Gröländischen Sprache, mit Theilweisem Einschluss des Labradordialects*. Berlin: G. Reimer.

Kneale, Willam, and Martha Kneale. 1962. *The Development of Logic*. Oxford: The Clarendon Press.

Kornfilt, Jaklin. 1997. *Turkish*. Descriptive Grammars. London: Routledge.

Kozinsky, I., V. Nedjalkov, and M. Polinskaja. 1988. Antipassive in Chukchee: oblique object, object incorporation, zero object. In Masayoshi Shibatani (ed.), *Passive and Voice*. Amsterdam: Benjamins, 651–706.

Kroeger, Paul. 1993. *Phrase-Structure and Grammatical Relations in Tagalog*. Stanford: CSLI.

Kroskrity, Paul. 1985. A holistic understanding of Arizona Tewa passives. *Language* 61, 306–28.

Kuno, Susumu. 1973. *The Structure of the Japanese Language*. Cambridge, MA: MIT Press.

1974. The position of relative clauses and conjunctions. *Linguistic Inquiry* 5, 117–36.

Lambrecht, K. 1994. *Information Structure and Sentence Form: Topic, Focus and the Mental Representation of Discourse Referents*. Cambridge: Cambridge University Press.

Langacker, Ronald W. 1976. *Non-distinct Arguments in Uto-Aztecan*. University of California Publications in Linguistics 82. Berkeley: University of California Press.

1987. Nouns and verbs. *Language* 63, 53–92.

Laughren, Mary. 1989. The configurationality parameter and Warlpiri. In Laszlo Karoly
    Maracz and Pieter Muysken (eds.), *Configurationality: The Typology of Asymme-*
    tries. Dordrecht: Foris, 319–53.
Lawler, John M. 1977. A agrees with B in Acehnese. In Peter Cole and Jerrold M.
    Sadock (eds.), *Syntax and Semantics*, vol. VIII: *Grammatical Relations.* New York:
    Academic Press, 219–48.
Lazdina, Tereza Budina. 1966. *Teach Yourself Latvian*. London: The English Universities
    Press.
Lee, Chung Min. 1973. *Abstract Syntax and Korean with Reference to English*. Seoul:
    Pan Korea Book Corp.
Lee, Kee-Dong. 1975. *Kusaiean Reference Grammar*. Honolulu: The University Press
    of Hawaii.
Lehmann, Winfred P. 1973. A structural principle of language and its implications.
    Language 49, 42–66.
    1978a. The great underlying ground-plans. In Winfred P. Lehmann (ed.), *Syntactic*
        Typology. Austin: University of Texas Press, 3–55.
    (ed.). 1978b. *Syntactic Typology*. Austin: University of Texas Press.
Levin, Beth, and Malka Rappaport Hovav. 1995. *Unaccusativity: At the Syntax – Lexical*
    Semantics Interface. Cambridge MA: MIT Press.
Levin, Lori. 1988. *Operations on Lexical Forms: Unaccusative Rules in Germanic*
    Languages. New York: Garland.
Levinson, Stephen. 1983. *Pragmatics*. Cambridge: Cambridge University Press.
Li, Charles N. (ed.). 1976. *Subject and Topic*. New York: Academic Press.
Li, Charles N., and Rainer Lang. 1979. The syntactic irrelevance of an ergative case
    in Enga and other Papuan languages. In Frans Planck (ed.), *Ergativity*. New York:
    Academic Press, 307–24.
Li, Charles N., and Sandra A. Thompson. 1976. Subject and topic: a new typology of
    language. In Charles N. Li (ed.), *Subject and Topic*. New York: Academic Press,
    457–89.
Li, Charles N., and Sandra A. Thompson. 1984. Mandarin. In William S. Chisholm *et al*.
    (eds.), *Interrogativity: A Colloquium on the Grammar, Typology and Pragmatics*
    of Questions in Seven Diverse Languages. Amsterdam: Benjamins, 47–61.
Li, Paul. 1973. *Rukai Structure*. Special Publication 64. Taipei: Institute of History and
    Philology, Academia Sinica.
Lohnstein, Horst. 2000. *Satzmodus – kompositionell. Zur Parametrisierung der Modus-*
    phrase im Deutschen. Berlin: Akademie-Verlag.
Louwerse, John. 1988. *The Morphosyntax of Una in Relation to Discourse Struc-*
    ture. Pacific Linguistics, Series B 100. Canberra: The Australian National
    University.
Lukas, Johannes. 1939. *Die Sprache der Buduma im Zentralen Sudan*. Abhandlungen für
    die Kunde des Morgenlandes, Band 24, nr. 2. Leipzig: Deutsche Morgenländische
    Gesellschaft. Reprinted 1966, Nendeln, Liechstenstein: Kraus Reprint Ltd.
Lyon, S. 1967. Tiahuitoltepec Mixe clause structure. *International Journal of American*
    Linguistics 23, 25–45.
Lyons, John. 1968. *Introduction to Theoretical Linguistics*. Cambridge: Cambridge
    University Press.
    1977. *Semantics*, 2 vols. Cambridge: Cambridge University Press.

Macaulay, Monica. 1996. *A Grammar of Chalcatongo Mixtec*. University of California Publications in Linguistics 127. Berkeley: University of California Press.

MacDonald, Lorna. 1994. The distribution of topics in Tauya discourse. In: G. Reesink (ed.), *Topics in Descriptive Papuan Linguistics*. Leiden: Research Group for the Languages and Cultures of Southeast Asia and Oceania, Leiden University, 75–97.

Mackay, Carolyn Joyce. 1999. *A Grammar of Misantla Totonac*. Salt Lake City: University of Utah Press.

Mahootian, Sharhzad. 1997. *Persian*. Descriptive Grammars. London: Routledge.

Manley, Timothy M. 1972. *Outline of the Sre Structure*. Honolulu: University of Hawaii Press.

Manning, Christopher D. 1996. *Ergativity: Argument Structure and Grammatical Relations*. Stanford: CSLI.

Marantz, Alec. 1984. *On the Nature of Grammatical Relations*. Cambridge, MA: MIT Press.

Masica, Colin. 1971. *Defining a Linguistic Area: South Asia*. Chicago: University of Chicago Press.

Matisoff, J. 1973. *The Grammar of Lahu*. Los Angeles: University of California Press.

Matthews, George H. 1965. *Hidatsa Syntax*. The Hague: Mouton.

Matthews, P. H. 1981. *Syntax*. Cambridge: Cambridge University Press.

Mazaudon, Martine. 1976a. La formation des propositions relatives en tibétain. *Bulletin de la Société de Linguistique de Paris* 73, 401–14.

Mazaudon, Martine. 1976b. Typological sketch of Tamang. Handout used in presentation at UCLA, 1976.

McGregor, William. 1990. *A Functional Grammar of Gooniyandi*. Amsterdam: John Benjamins.

McLendon, Sally. 1975. *A Grammar of Eastern Pomo*. Berkeley: University of California Press.

Merlan, Francesca. 1982. *Mangarayi*. Lingua Descriptive Studies 4. Amsterdam: North-Holland.

   1983. *Ngalakan Grammar, Texts, and Vocabulary*. Pacific Linguistics, Series B 89. Canberra: The Australian National University.

   1985. Split intransitivity: functional oppositions in intransitive reflection. In Johanna Nichols and Anthony Woodbery (eds.), *Grammar Inside and Outside the Clause*. Cambridge: Cambridge University Press, 324–62.

Meyer, Roland, and Gerhild Zybatow. 2003. Zur Satztypologie russischer Entscheidungsfragen: Satztypfestigung durch Intonation? Paper presented at the 25th meeting of the Deutsche Gesellschaft für Sprachwissenschaft. Munich.

Michaelis, Laura. 2001. Exclamative constructions. In Martin Haspelmath, Ekkehard König, Wulf Oesterreicher, and Wolfgang Raible (eds.), *Language Typology and Language Universals*. Berlin: de Gruyter, 1038–58.

Michailovsky, Boyd. 1988. *La langue hayu*. Paris: Editions du CNRS.

Mirikitani, Leatrice T. 1972. *Kapampangan Syntax*. Oceanic Linguistics Special Publication 10 Honolulu: The University Press of Hawaii.

Mithun, Marianne. 1991. Active/agentive case marking and its motivations. *Language* 67, 510–46.

   1992. Is basic word order universal? In Doris Payne (ed.), *Pragmatics of Word Order Flexibility*. Amsterdam: Benjamins, 15–62.

Mohanan, K. P. 1982. Grammatical relations and clause structure in Malayalam. In Joan Bresnan (ed.), *The Mental Representation of Grammatical Relations*. Cambridge, MA: MIT Press, 504–89.

Mohanan, T. 1994. *Argument Structure in Hindi*. Stanford: CSLI.

Mosel, Ulrike. 1984. *Tolai Syntax and its Historical Development*. Pacific Linguistics, Series B 92. Canberra: The Australian National University.

Mosel, Ulrike, and Even Hovdhaugen. 1992. *Samoan Reference Grammar*. Oslo: Scandinavian University Press.

Moyse-Faurie, Claire. *Grammaire du Futunien*. Collection Université. Nouméa, New Caledonia: Coédition Centre de documentation pédagogique Nouvelle-Calédonie, Association socio-culturelle de Futuna, Service des Affaires culturelles de Futuna.

Munro, Pamela, and Lynn Gordon. 1982. Syntactic relations in Western Muskogean. *Language* 58, 81–115.

Naish, Constance. 1979. *A Syntactic Study of Tlingit*. Dallas: Summer Institute of Linguistics.

Nash, David George. 1980. *Topics in Warlpiri Grammar*. New York: Garland.

Naylor, L. B. 1925. *A Practical Handbook of the Chin Language (Siyin Dialect): Containing Grammatical Principles with Numerous Exercises and a Vocabulary*. Rangoon: Superintendent, Union Government Printing and Stationery.

Nedjalkov, Igor. 1997. *Evenki*. Descriptive Grammars. London: Routledge.

Newell, Leonard E. 1993. *Batad Ifugao Dictionary: With Ethnographic Notes*. Manila: Linguistic Society of the Philippines.

Newman, Paul. 1970. *A Grammar of Tera*. University of California Publications in Linguistics 57. Berkeley: University of California Press.

  1971. The Hausa negative markers. *Studies in African Linguistics* 2, 183–95.

Nguyen, Phu Phong. 1976. *Le syntagme verbal en Vietnamien*. Paris: Mouton.

Nichols, Johanna. 1979. Syntax and pragmatics in Manchu-Tungus languages. In P. Clyne, W. Hanks, and C. Hofbauer (eds.), *The Elements: A Parasession on Linguistic Units and Levels Including Papers from the Conference on Non-Slavic Languages of the USSR*. Chicago: Chicago Linguistics Society, 420–8.

  1986. Head-marking and dependent-marking grammar. *Language* 62, 56–119.

  1994. Ingush. In R. Smeets (ed.), *The Indigenous Languages of the Caucasus*. Delmar, NY: Caravan Books, 79–146.

Nichols, Johanna, and Anthony C. Woodbury (eds.). 1985. *Grammar Inside and Outside the Clause*. Cambridge: Cambridge University Press.

Noonan, Michael P. 1992. *A Grammar of Lango*. Mouton Grammar Library 7. Berlin: Mouton de Gruyter.

  1994. A tale of two passives in Irish. In Barbara Fox and Paul Hopper (ed.), *Voice: Form and Function*. Amsterdam: John Benjamins, 271–311.

Noonan, Michael P., and Edith Bavin-Woock. 1978. The passive analog in Lango. *Proceedings of the Berkeley Linguistics Society* 4, 128–39.

Nordlinger, Rachel. 1998. *A Grammar of Wambaya, Northern Territory (Australia)*. Pacific Linguistics, Series C 140. Canberra: Australian National University.

Oates, Lynette F. 1988. *The Muruwari Language*. Pacific Linguistics, Series C, no. 108. Canberra: Australian National University.

Oehrle, Richard. 1976. The grammatical status of the English dative alternation. Doctoral dissertation. MIT Cambridge, MA.

Ohala, John J. 1983. Cross-language use of pitch: an ethological view. *Phonetica* 40, 1–18.

1994. The frequency code underlies the sound-symbolic use of voice pitch. In Leanne Hinton, Johanna Nichols, and John J. Ohala (eds.), *Sound Symbolism*. Cambridge: Cambridge University Press, 325–47.

Önnerfors, Olaf. 1997. *Verb-erst-Deklarativsätze*. Stockholm: Almquist and Wiksel International.

Osborne, C. R. 1974. *The Tiwi Language*. Canberra: Australian Institute of Aboriginal Studies.

Owens, Jonathan. 1985. *A Grammar of Harar Oromo (Northeastern Ethiopia)*. Hamburg: Buske.

Palmer, F. R. 1986. *Mood and Modality*. Cambridge: Cambridge University Press.

1994. *Grammatical Roles and Relations*. Cambridge: Cambridge University Press.

Pawley, Andrew. 1993. A language which defies description by ordinary means. In W. A. Foley (ed.), *The Role of Theory in Language Description*. Berlin: Mouton de Gruyter, 87–130.

Payne, Doris L. 1987. Information structuring in Papago narrative discourse. *Language* 63, 783–804.

1990. *The Pragmatics of Word Order: Typological Dimensions of Verb-initial Languages*. Berlin: Mouton de Gruyter.

1994. The Tupí-Guaraní inverse. In Barbara Fox and Paul J. Hopper (eds.), *Voice: Form and Function*. Amsterdam: Benjamins, 313–40.

Payne, Doris L., and Thomas E. Payne. 1990. Yagua. In Desmond C. Derbyshire and Geoffrey K. Pullum (eds.), *Handbook of Amazonian Languages*, vol. II. Berlin: Mouton de Gruyter, 249–474.

Payne, Thomas E. 1982. Role and reference related subject properties and ergativity in Yup'ik Eskimo and Tagalog. *Studies in Language* 6, 75–196.

Perlmutter, David M. 1978. Impersonal passives and the unaccusative hypothesis. *Berkeley Linguistics Society* 4, 157–89.

(ed.). 1983. *Studies in Relational Grammar 1*. Chicago: University of Chicago Press.

Perlmutter, David, and Paul Postal. 1983. Towards a universal characterization of passive. In David Perlmutter (ed.), *Studies in Relational Grammar*, vol. I. Chicago: University of Chicago Press, 3–29.

Perlmutter, David, and Richard Rhodes. 1988. Thematic-syntactic alignments in Ojibwa: evidence for subject–object reversal. Paper delivered at the LSA, New Orleans.

Pickett, Velma. 1960. *The Grammatical Hierarchy of Isthmus Zapotec*. Language Dissertation 56. Baltimore: *Linguistic Society of America*.

Pinker, Steven. 1989. *Learnability and Cognition: The Acquisition of Argument Structure*. Cambridge, MA: MIT Press.

Plank, Frans. 1984. *Objects: Towards a Theory of Grammatical Relations*. New York: Academic Press.

(ed.), 1995. *Double Case: Agreement by Suffixaufnahme*. Oxford: Oxford University Press.

Polinsky, Maria. 2005. Antipassive constructions. In Haspelmath *et al.* (eds.), *The World Atlas of Language Structures*. Oxford: Oxford University Press, 438–41.

Pope, Emily N. 1976. *Questions and Answers in English*. The Hague: Mouton.

Postal, Paul. 1986. *Studies in Passive Clauses*. Albany: State University of New York Press.

Prasad, Bal Ram. 1991. *Mising Grammar*. Mysore, India: Central Institute of Indian Languages.

Prince, Ellen F. 1978. A comparison of WH-clefts and IT-clefts in discourse. *Language* 54, 883–906.

  1981. Topicalization, focus-movement and Yiddish-movement: a pragmatic differentiation. *Proceedings of the Berkeley Linguistics Society* 7, 249–64.

Prost, R. P. A. 1956. *La langue Sonay et ses dialectes*. Mémoires de l'Institut Français d'Afrique Noire 47. Dakar: Ifan.

Rajaona, Simeon. 1972. *Structure du Malgache: étude des formes prédicatives*. Fianarntsoa, Madagascar: Ambozontany.

Randriamasimanana, Charles. 1986. *The Causatives of Malagasy*. Honolulu: University of Hawaii Press.

Rangan, K. 1979. *Purki Grammar*. Manasagangotri, Mysore: Central Institute of Indian Languages.

Raz, Shlomo. 1983. *Tigre Grammar and Texts*. Malibu: Undena.

Reis, Marga. 1992. Zur Grammatik und Pragmatik von Echo-w-Fragen. In Inger Rosengren (ed.), *Satz und Illokution*. Linguistische Arbeiten 278. Tübingen: Niemeyer, vol. I, 213–62.

  1999. On sentence types in German: an enquiry into the relationship between grammar and pragmatics. *Interdisciplinary Journal for Germanic Linguistics and Semiotic Analysis* 4, 195–236.

Rennison, John R. 1997. *Koromfe*. London: Routledge.

Rhodes, Richard A. 1976. The morphosyntax of the Central Ojibwa verb. Dissertation. University of Michigan.

  1990. Ojibwa secondary objects. In Katarzyna Dziwirek, Patrick Farrel, and Errapel Majías Bikandi (eds.), *Grammatical Relations: A Cross-Theoretical Perspective*. Stanford: CSLI, 401–14.

Rice, Keren. 1989. *A Grammar of Slave*. Berlin: Mouton de Gruyter.

Robins, R. H. 1964. *General Linguistics: An Introductory Survey*. London: Longmans, Green and Co.

Robinson, Stuart. 2002. Constituent order in Tenejapa Tzeltal. *International Journal of American Linguistics* 68, 51–81.

Rochemont, M., and P. Culicover. 1990. *English Focus Constructions and the Theory of Grammar*. Cambridge: Cambridge University Press.

Rögnvaldsson, Eiríkur. 1982. We need (some kind of a) rule of Conjunction Reduction. *Linguistic Inquiry* 13, 557–61.

Roland, Katy. 1994. The pragmatics of Modern Greek voice: active, inverse, and passive. In T. Givón (ed.), *Voice and Inversion*. Amsterdam: John Benjamins, 233–60.

Rosengren, Inger (ed.). 1992a. *Satz und Illokution*, vol. I. Linguistische Arbeiten 278. Tübingen: Niemeyer.

  1992b. Zur Grammatik und Pragmatik der Exklamation. In Inger Rosengren (ed.), *Satz und Illokution*, vol. I. Linguistische Arbeiten 278. Tübingen: Niemeyer, 263–306.

  (ed.). 1993. *Satz und Illokution*, vol. II. Linguistische Arbeiten 278 Tübingen: Niemeyer.

Ross, Malcolm, and John Natu Paol. 1978. *A Waskia Grammar Sketch and Vocabulary*. Pacific Linguistics, Series B 56. Canberra: Australian National University.

Rubino, Carl R. Galvez. 1998. *Ilocano: Ilocano–English/English–Ilocano Dictionary and Phrasebook*. New York: Hippocrene Books.

Rupp, James E. 1989. *Lealao Chinantec Syntax*. Studies in Chinantec Languages 2. Dallas: Summer Institute of Linguistics and the University of Texas at Arlington.

Sadock, Jerrold M. 1974. *Toward a Linguistic Theory of Speech Acts*. New York: Academic Press.

    1984. West Greenlandic. In William S. Chisholm *et al*. (eds.), *Interrogativity: A Colloquium on the Grammar, Typology and Pragmatics of Questions in Seven Diverse Languages*. Amsterdam: Benjamins.

    1991. *Autolexical Syntax: A Theory of Parallel Grammatical Representations*. Chicago: University of Chicago Press.

    1994. Toward a grammatically realistic typology of speech acts. In Svas L. Tsohatzidis (ed.), *Foundations of Speech Acts Theory. Philosophical and Linguistics Perspectives*. London: Routledge, 393–406.

Sadock, Jerrold M., and Arnold M. Zwicky. 1985. Speech act distinctions in syntax. In T. Shopen (ed.), *Language Typology and Syntactic Description*, vol I. Cambridge: Cambridge University Press, 155–96.

Saltarelli, Mario. 1988. *Basque*. Croom Helm Descriptive Grammars. London: Croom Helm.

Sapir, Edward. 1921. *Language*. New York: Harcourt Brace.

    1930. Southern Paiute, a Shoshonean language. *Proceedings of the American Academy of Arts and Sciences* 65, 1–296.

Sapir, Edward, and Morris Swadesh. 1960. *Yana Dictionary*. Berkeley: University of California Press.

Sbisà, Marina. 1995. Speech act theory. In Jef Verschueren and Jan-Ola Östman, and Jan Blommaert (eds.), *Handbook of Pragmatics*. Amsterdam: Benjamins, 495–505.

Scancarelli, Janine. 1986. Pragmatic roles in Cherokee Grammar. *Proceedings of the Twelfth Annual Meeting of the Berkeley Linguistics Society*, 224–34.

Schachter, Paul. 1974. A non-transformational account of serial verbs. *Studies in African Linguistics*. Supplement 5, 253–70.

    1976. The subject in Philippine languages: topic, actor, actor-topic or none of the above. In Charles N. Li (ed.), *Subject and Topic*. New York: Academic Press, 491–518.

    1977. Reference-related and role-related properties of subjects. In Peter Cole and Jerrold M. Sadock (eds.), *Syntax and Semantics*, vol. VIII: *Grammatical Relations*. New York: Academic Press, 279–306.

    1978. English propreciates. *Linguistic Analysis* 4, 187–224.

    1996. *The subject in Tagalog: still none of the above*. UCLA Occasional Papers in Linguistics 15. Los Angeles: University of California, Los Angeles.

Schachter, Paul, and Fe T. Otanes. 1972. *Tagalog Reference Grammar*. Berkeley: University of California Press.

Schaub, Willi. 1985. *Babungo*. Croom Helm Descriptive Grammars. Dover, NH: Croom Helm.

Schwartz, Arthur. 1976. On the universality of Subjects. In Charles N. Li (ed.), *Subject and Topic*. New York: Academic Press, 516–43.

Searle, John. 1969. *Speech Acts*. Cambridge: Cambridge University Press.

  1976. The classification of illocutionary acts. *Language in Society* 5, 1–24.

Searle, John, Ferenc Kiefer, and Manfred Bierwisch. 1980. *Speech Act Theory and Pragmatics*. Dordrecht: Reidel.

Seidenadel, Carl Wilhelm. 1909. *The First Grammar of the Language Spoken by the Bontoc Igorot with Vocabulary and Texts*. Chicago: Open Court Publishing Co.

Seiler, Walter. 1985. *Imonda, a Papuan Language*. Pacific Linguistics, Series B 93. Canberra: The Australian National University.

Shibatani, Masayoshi (ed.). 1988. *Passive and Voice*. Amsterdam: Benjamins.

Shopen, Timothy (ed.). 1987a [1979]. *Languages and their Speakers*. Philadelphia: University of Pennsylvania Press.

  (ed.). 1987b. *Languages and their Status*. Philadelphia: University of Pennsylvania Press.

Siemund, Peter. 2001. Interrogative constructions. In Martin Haspelmath, Ekkehard König, Wulf Oesterreicher, and Wolfgang Raible (eds.), *Language Typology and Language Universals*. Berlin: de Gruyter, 1010–28.

Siewierska, Anna. 1984. *The Passive: A Comparative Linguistic Analysis*. London: Croom Helm.

  2005a. Alignment of verbal person marking. In M. Haspelmath *et al*. (eds.), *The World Atlas of Language Structures*. Oxford: Oxford University Press, 406–9.

  2005b. Passive constructions. In M. Haspelmath *et al*. (eds.), *The World Atlas of Language Structures*. Oxford: Oxford University Press, 434–7.

Silverstein, Michael. 1976. Hierarchy of features and ergativity. In R. M. W. Dixon (ed.), *Grammatical Categories in Australian Languages*. Canberra: Australian Institute of Aboriginal Studies, 112–71.

Silzer, Peter. 1983. Ambai: an Austronesian language of Irian Jaya, Indonesia. Unpublished Ph.D. thesis. The Australian National University.

Simpson, Jane H. 1991. *Warlpiri Morpho-syntax*. Dordrecht: Kluwer Academic.

Snyman, J. W. 1970. *An Introduction to the !Xu (!Kung) Language.* Capetown: A. A. Balkema.

Sridhar, S. N. 1990. *Kannada.* Descriptive Grammars. London: Routledge.

Staalsen, Philip 1972. Clause relationships in Iatmul. In R. K. Lewis *et al*. (eds.), *Papers in New Guinea Linguistics, 15*. Pacific Linguistics series A 31. Canberra: The Australian National University, 45–68.

Stassen, Leon. 1997. *Intransitive Predication*. Oxford: Clarendon Press.

  2005a. Predicate adjectives. In M. Haspelmath *et al*. (eds.), *The World Atlas of Language Structures*. Oxford: Oxford University Press, 478–81.

  2005b. Nominal and locational predication. In M. Haspelmath *et al*. (eds.), *The World Atlas of Language Structures*. Oxford: Oxford University Press, 482–5.

Stevenson, R. C. 1969. *Bagirmi Grammar*. Linguistic Monograph Series 3. Khartoum: Sudan Research Unit, University of Khartoum.

Stirling, Lesley F. 1996. *Switch-Reference and Discourse Representation*. Studies in Linguistics Series. Cambridge: Cambridge University Press.

Story, Gillian. 1979. *A Morphological Study of Tlingit*. Dallas: Summer Institute of Linguistics.

Sugita, H. 1973. Semitransitive verbs and object incorporation in Micronesian languages. *Oceanic Linguistics* 12, 393–406.

Sulkala, Helena, and Merja Karjalainen. 1992. *Finnish*. London: Routledge.

Svartvik, Jan. 1966. *On Voice and the English Verb*. The Hague: Mouton.

Swadesh, Morris. 1939. Nootka internal syntax. *International Journal of American Linguistics* 9, 77–102.

Swartz, Stephen (ed.). 1982. *Papers in Warlpiri Grammar: In Memory of Lothar Jagst. Work-papers of SIL-AAB*, Series A 6. Berrimah, Australia: Summer Institute of Linguistics.

Taylor, Charles. 1985. *Nkore-Kiga*. Croom Helm Descriptive Grammars. Dover, NH: Croom Helm.

Thomas, David D. 1971. *Chrau Grammar*. Honolulu: University of Hawaii Press.

Thompson, Chad. 1990. On the treatment of topical objects in Chepang: passive or inverse. *Studies in Language* 14, 405–27.

Thompson, E. D. 1983. Kunama: phonology and noun phrase. In M. L. Bender (ed.), *Nilo-Saharan Language Studies*. East Lansing: African Studies Center, Michigan State University, 280–322.

Thompson, Laurence C. 1965. *A Vietnamese Grammar*. Seattle: University of Washington Press.

Thráinsson, Höskuldur. 1979. *On Complementation in Icelandic*. New York: Garland.

Timberlake, Alan. 1976. Subject properties of the North Russian passive. In Charles N. Li (ed.), *Subject and Topic*. New York: Academic Press, 545–57.

Traugott, Elizabeth C. 1985. On conditional. In John Haiman (ed.), *Iconicity in Syntax*. Amsterdam: Benjamins, 289–307.

Trithart, L. 1977. *Relational Grammar and Chichewa Subjectivization Rules*. Bloomington: Indiana University Linguistics Club.

Tryon, Darrell T. 1967. *Nengone Grammar*. Pacific Linguistics, Series B 6. Canberra: The Australian National University.

   1970a. *Conversational Tahitian*. Berkeley: University of California Press.

   1970b. *An Introduction to Maranungku*. Pacific Linguistics, Series B 15. Canberra: The Australian National University.

Tsohatzidis, Svas L. (ed.). 1994. *Foundations of Speech Acts Theory. Philosophical and Linguistics Perspectives*. London: Routledge.

Tsunoda, Tasaku. 1981. *The Djaru Language of Kimberley, Western Australia*. Pacific Linguistics, Series B 78. Canberra: The Australian National University.

Tucker, Archibald N. 1940/1967. *The Eastern Sudanic Languages,* vol. I. London: Dawsons of Pall Mall, for the International African Institute.

Tuggy, David H. 1977. Tetelcingo Nahuatl. In Ronald W. Langacker (ed.), *Modern Aztec Grammatical Sketches*. Studies in Uto-Aztecan Grammar 2. Arlington: Summer Institute of Linguistics, 1–140.

Ultan, Russell. 1978. Some general characteristics of interrogative systems. In Joseph H. Greenberg (ed.), *Universals of Human Language*. Stanford: Stanford University Press, 211–48.

Upadhyaya, U. P., and N. D. Krishnamurthy. 1972. *Conversational Kannada*. Dharwar, India: Lakshmibai Gurunath Joshi.

Van den Berg, Helma. 1995. *A Grammar of Hunzib*. Munich: Lincom Europa.

Van den Berg, René. 1989. *A Grammar of the Muna Language*. Dordrecht: Foris.

van der Auwera, Johan, and Ludo Lejeune. 2005a. The morphological imperative. In Martin Haspelmath, Matthew Dryer, David Gil, and Bernard Comrie (eds.), *World Atlas of Language Structures*. Oxford: Oxford University Press.

2005b. The prohibitive. In Martin Haspelmath, Matthew Dryer, David Gil, and Bernard Comrie (eds.), *World Atlas of Language Structures*. Oxford: Oxford University Press.

van der Auwera, Johan, Nina Dobrushina, and Valentin Goussev. 2004. A semantic map for imperative–hortatives. In D. Willems, T. Colleman, and B. Defrancq (eds.), *Contrastive Analysis of Language: Identifying Linguistic Units of Comparison*. Basingstoke: Palgrave Macmillan, 44–66.

2005. Imperative–hortative systems. In Martin Haspelmath, Matthew Dryer, David Gil, and Bernard Comrie (eds.), *World Atlas of Language Structures*. Oxford: Oxford University Press.

Van Enk, Gerrit J., and Lourens De Vries. 1997. *The Korowai of Irian Jaya: Their Language in its Cultural Context*. New York: Oxford University Press.

Van Klinken, Catharina Lumien. 1999. *A Grammar of the Fehian Dialect of Tetun*. Pacific Linguistics, Series C 155. Canberra: The Australian National University.

Van Valin, Robert D. 1981. Grammatical relations in ergative languages. *Studies in Language* 5, 361–94.

1985. Case marking and the structure of the Lakhota clause. In Johanna Nichols and Anthony Woodbury (eds.), *Grammar Inside and Outside the Clause*. Cambridge: Cambridge University Press, 363–413.

1990. Semantic parameters of split intransitivity. *Language* 66, 221–60.

Van Valin, Robert D., and Randy J. LaPolla. 1997. *Syntax: Structure, Meaning and Function*. Cambridge: Cambridge University Press.

Vennemann, Theo. 1973. Explanation in syntax. In John Kimball (ed.), *Syntax and Semantics 2*. New York: Seminar Press, 1–50.

1974a. Analogy in generative grammar: the origin of word order. In Luigi Heilmann (ed.), *Proceedings of the Eleventh International Congress of Linguists (1972)*. Bologna: Il Mulino, 79–83.

1974b. Theoretical word order studies: results and problems. *Papiere zur Linguistik* 7, 5–25.

1976. Categorial grammar and the order of meaningful elements. In Alphonse Juilland (ed.), *Linguistic Studies Offered to Joseph Greenberg on the Occasion of His Sixtieth Birthday*. Saratoga, CA: Anma Libri, 615–34.

Vennemann, Theo, and Ray Harlow. 1977. Categorial grammar and consistent basic VX serialization. *Theoretical Linguistics* 4, 227–54.

Verma, Manindra K., and K. P. Mohanan (eds.). 1990. *Experiencer Subjects in South Asian Languages*. Stanford: CSLI.

Vogel, Petra M., and Bernard Comrie (eds.). 2000. *Approaches to the Typology of Word Classes*. Berlin: Mouton de Gruyter.

Walsh, Michael. 1976. The Mujinypata language. Ph.D. thesis. Australian National University.

Walters, David E. 2002. *A Grammar of Kham*. Cambridge: Cambridge University Press.

Warotamasikkhadit, U. 1972. *Thai Syntax*. The Hague: Mouton.

Weber, David J. 1976. *Suffix-as-operator Analysis and the Grammar of Successive Encoding in Llacon (Huanuco) Quechua*. Yarinacocha, Peru: Instituto Lingüístico de Verano.

  1989. *A Grammar of Huallaga (Huánaco) Quechua*. University of California Publications in Linguistics 112. Berkeley: University of California Press.

Wechsler, Stephen. 1995. *The Semantic Basis of Argument Structure*. Stanford: CSLI.

Wechsler, Stephen, and I. Wayan Arka. 1998. Syntactic ergativity in Balinese: an argument-structure based theory. *Natural Language and Linguistic Theory* 16, 384–441.

Wells, Margaret A. 1979. *Siroi Grammar*. Pacific Linguistics, Series B 51. Canberra: The Australian National University.

Welmers, W. E., and B. F. Welmers. 1969. Noun modifiers in Igbo. *International Journal of American Linguistics* 35, 315–22.

Whistler, Kenneth. 1985. Focus, perspective, and inverse person marking in Nootkan. In Johanna Nichols and Anthony C. Woodbury (eds.), *Grammar Inside and Outside the Clause*. Cambridge: Cambridge University Press, 227–65.

Wickman, B. 1955. *The form of the object in Uralic languages*. Uppsala Universitets Arsskrift 6. Uppsala: Almqvist & Wiksells.

Wierzbicka, Anna. 1986. What's in a noun? (or: how do nouns differ in meaning from adjectives?). *Studies in Language* 10, 353–89.

  1987. *English Speech Act Verbs*. Sydney: Academic Press.

  2000. Lexical prototypes as a universal basis for cross-linguistic identification of parts of speech. In P. M. Vogel and B. Comrie (eds.), *Approaches to the Typology of Word Classes*. Berlin: Mouton de Gruyter, 285–317.

Winstedt, R. O. 1914. *Malay Grammar*. Oxford: Clarendon Press.

Witherspoon, G. 1980. Language in culture and culture in language. *International Journal of American Linguistics* 46, 1–13.

Wolfart, H. Christoph. 1973. *Plains Cree: A Grammatical Study*. Transactions of the American Philosophical Society, New Series, Vol. 63, Part 5. Philadelphia: American Philosophical Society.

Wolfart, H. Christoph, and Janet Carroll. 1981. *Meet Cree*, 2nd edn. Edmonton: University of Alberta Press.

Wolff, John V. 1966. *Beginning Cebuano, Part 1*. New Haven: Yale University Press.

Wunderlich, Dieter. 1976. *Studien zur Sprechakttheorie*. Frankfurt am Main: Suhrkamp.

  1991. Bedeutung und Gebrauch. In Arnim von Stechow and Dieter Wunderlich (eds.), *Semantik. Ein internationales Handbuch zeitgenössischer Forschung*. Berlin: de Gruyter, 32–52.

Xrakovskij, Viktor S. 2001. Hortative constructions. In Martin Haspelmath, Ekkehard König, Wulf Oesterreicher, and Wolfgang Raible (eds.), *Language Typology and Language Universals*. Berlin: de Gruyter, 1028–38.

Zaefferer, Dietmar. 1990. On the coding of sentential modality. In Johannes Bechert, Giuliano Bernini, and Claude Buridant (eds.), *Towards a Typology of European Languages*. Berlin: Mouton de Gruyter, 215–37.

Zaenen, Annie. 1983. On syntactic binding. *Linguistic Inquiry* 14, 469–504.
   1993. Unaccusativity in Dutch: integrating syntax and lexical semantics. In James
      Pustejovsky (ed.), *Semantics and the Lexicon*. Dordrecht: Kluwer, 129–62.
Zanuttini, Raffaella, and Paul Portner. 2003. Exclamative clauses: at the syntax–
      semantics interface. *Language* 79, 1, 39–81.

Subject index