5

SUBJECT

5.1 THE PROBLEM

In section 3.3, we mentioned some of the problems inherent in working with grammatical relations, including subject, and some of the possible approaches towards a solution to these problems. In the present chapter, we will look in considerably more detail at one particular aspect of this problem, namely the definition of subject cross-linguistically. Subject is an important notion, used frequently, both in traditional grammar and in more recent linguistic work, both in the descriptions of individual languages and in stating cross-linguistic generalizations. If linguists were invariably in agreement in stating which noun phrase, in each construction in each language, is the subject, then we could, perhaps, accept this inter-subjective agreement, and devote correspondingly less energy to trying to find an explicit definition of subject. However, it turns out that, in a wide range of cases, this inter-subjective agreement is lacking, so that the need does arise as a serious empirical problem to establish criteria for declaring a given noun phrase to be or not to be a subject.

One particular instance of lack of agreement among linguists on subjecthood is illustrated by competing analyses of the ergative construction. We shall return below, in sections 5.3 and 6.2.2 to a more detailed discussion of ergativity, and for present purposes we may simply give some illustrative examples of the kind of problem that arises, using Chukchi as our example:

ɣəm tə-yet-ɣʔek. (1)
I-ABSOLUTIVE came-1SINGULAR
'I came.'

ɣəm-nan ɣət tə-lʔu-ɣət. (2)
I-ERGATIVE thou-ABSOLUTIVE saw-1SINGULAR-2SINGULAR
'I saw thee.'

Analyses of English agree that, in the English versions of these two sentences, *I* is subject both of the intransitive construction of (1) and of the transitive construction of (2); moreover, English morphology, at least for pronouns, exactly mirrors this distribution: the subjects are in the nominative, the direct object in the accusative. In Chukchi, as in English, there are two cases used for these three noun phrases, but their distribution is quite different: the absolutive case is used to translate *I* (intransitive subject) of (1), and to translate *thee* (direct object) of (2), whereas a separate case, the ergative, is used to translate *I* (transitive subject) of (2). The question therefore arises whether, in Chukchi, one should not rather group together the absolutive noun phrases as subject, following the morphology, rather than simply following the distribution that turns out to be relevant for English. Although in early periods many linguists working on ergative languages tried to solve this problem a priori, by fiat – and in either direction, by relying on the morphology or by disregarding it – the question is in fact an empirical question, and in sections 5.3–4 we will see that its answer is much less simple than either of these solutions. For the moment, however, we may simply note that the problem exists.

Of course, in addition to criteria of case marking in establishing subjecthood, it will be clear from the discussion of section 3.3 that syntactic criteria are also important in establishing subjecthood. In English, for instance, we can note the following two syntactic criteria of subjecthood. First, verbs agree in person and number with their subject; although English verb morphology is fairly atrophied, this distinction is still maintained consistently in the difference in the present tense between third person singular and all other forms, and in a few other instances with irregular verbs, so that we have the third person singular form in *he sees you* but the non-third person singular form in *I see you*. Secondly, in the kinds of constructions called subject-to-object raising by many transformational-generative grammarians, we find that the subject of a *that*-clause, and only the subject, can, after certain verbs, appear in an alternative construction of type (4):

I believe that Max is an accountant. (3)

I believe Max to be an accountant. (4)

In the vast majority of sentence-types, these two syntactic criteria coincide, i.e. there is agreement between logically independent criteria as to the subject in English. There are, however, some sentence types where this agreement is not found, such as sentences introduced by *there is/are*:

There are unicorns in the garden. (5)

There is a unicorn in the garden. (6)

In such examples, at least in the standard language, verb agreement is determined by the noun phrase that follows *there is/are*. Subject-to-object raising, however, treats *there* as the subject, giving:

I believe there to be a unicorn/unicorns in the garden. (7)

And, indeed, in such instances we find disagreement as to which noun phrase, in (5) and (6), should be considered the subject: different weighting of different criteria gives different results. So even in English there are some construction types where there is no agreement among linguists as to which noun phrase is subject.

Faced with such problems surrounding the characterization of the notion subject, there are two possible approaches. On the one hand, one could claim that the notion of subject is misleading from the outset, and should be banished from linguistic theory. On the other hand, one could try and work out a definition of subjecthood which, while corresponding to linguists' inter-subjective intuitions in the clear cases, would also make insightful claims about the unclear cases. In the present chapter, we follow the second of these paths. Before embarking on the details of the definition, however, we should make some further preliminary remarks. First, we are not committed a priori to the view that subject is a necessary descriptive category in the grammar of every language: there may well be languages where it is not appropriate, though equally there are languages (including English) where it is appropriate. Secondly, we are not committed to the view that, even in a language where subject is generally valid, every sentence will necessarily have a subject. Thirdly, we are not committed to the view that the translation of a sentence from language X where a certain noun phrase is subject will necessarily have that same noun phrase as subject in language Y. Examples of all of these points will occur below.

Finally, although we will argue that the notions of topic and agent must play a role in the definition of subject, we argue that, even in English, it is clear that the notion of subject cannot be identified with either of these notions. If we take, for instance, our criterion of verb-agreement, then it is clear that in the passive sentence *the men were hit by the boy*, the plural verb *were* does not agree with the agent; and it is equally clear that in the topicalized sentence *John I know* the non-third person singular verb is not in agreement with the topic. However close the connection may be between grammatical relations, semantic roles, and pragmatic roles, they cannot be identified with one another.

5.2 ON DEFINITIONS AND CATEGORIES

Before turning specifically to the definition of subject, it is necessary for us to make some preliminary remarks on the nature of definitions, in particular on the nature of definitions of linguistic categories, in order to avoid certain later misunderstandings. The kind of definition of subject towards which we will be working is the following: the prototype of subject represents the intersection of agent and topic, i.e. the clearest instances of subjects, cross-linguistically, are agents which are also topics. There are two important characteristics of this definition: first, it is multi-factor; second, it is stated in terms of prototypes, rather than in terms of necessary and sufficient criteria for the identification of subjects. The second point is particularly important, given that many subjects in many constructions in many languages are not topic, or are not agent, or are neither.

The use of a multi-factor definition is unlikely to raise any eyebrows, since such definitions are quite widespread in linguistics and other areas, as for instance if we define preposition in terms of the intersection of adposition and position in front of the governed noun phrase. However, the attempt to use definitions in terms of prototypes for linguistic categories has met with an inordinate amount of opposition and prejudice, so that it is worth spending some time on discussion of this issue. Rather than discussing the problem directly in terms of subject properties, we will use some other examples, where the use of prototypes is much more clearly justified. Note that the use of these analogies does not in itself justify the use of a prototype-based definition of subject, but it does demonstrate that we cannot a priori reject this kind of definition, but must rather weigh up the pros and cons in terms of their fit with the data and their evaluation relative to alternative definitions.

In chapter 2, we illustrated one very clear area where definitions of categories in terms of prototype seem to be required, namely with colour terms, where humans seem to recognize a central, focal value for a colour term, rather than clear-cut boundaries. What this means is that there is no set of necessary and sufficient conditions that an object must satisfy in order to be called, for instance, red. But equally, this does not mean that we can state no restrictions on the use of the term red: this term is most appropriate for the focal value, and less and less appropriate as one moves away from this focal area and approaches the foci of other colour terms. This example thus establishes that there is at least one area where humans do categorize in terms of prototypes, thus opening up this kind of definition as a real possibility.

Similar examples can also be found using more clearly linguistic categories, and the example we will use here concerns the distinction between nouns and adjectives in Russian, in particular the relation of numerals to these two. In Russian, in general, the distinction between nouns and adjectives is clear-cut, so that we can establish criteria that correlate with the focal values (prototypes) of noun and adjective. Numerals, however, fall in between these two prototypes, in a way that makes impossible any estab-

lishment of non-arbitrary cut-off points. In distinguishing adjectives from nouns, we may take two comparable construction types, the first being a noun phrase consisting of an attributive adjective and head noun (e.g. *xorošij mal'čik* 'good boy'), the second being a quantity phrase consisting of a head noun defining the quantity and a dependent genitive defining the entity being measured (e.g. *stado ovec* 'flock of sheep').

The following criteria characterize the adjective in the attributive construction: (a) the adjective agrees in gender with its head noun, following a three-way masculine/feminine/neuter distinction (though only in the singular), e.g. *xorošij mal'čik* 'good boy', *xorošaja devočka* 'good girl', *xorošeje okno* 'good window'; (b) the adjective agrees in number with its head noun, on a singular/plural opposition, e.g. *xorošij mal'čik* 'good boy', *xorošije mal'čiki* 'good boys'; (c) the adjective agrees in case with its head noun, e.g. nominative *xorošij mal'čik*, but dative *xorošemu mal'čiku*, instrumental *xorošim mal'čikom*; (d) many nouns have distinct accusative forms depending on whether or not they are animate, and adjectives agree with their head noun in terms of this distinction, e.g. inanimate accusative *xorošij stol* 'good table', animate accusative *xorošego mal'čika* 'good boy', even though both *stol* and *mal'čik* are masculine singular. Head nouns in the quantitative construction have none of these properties. Thus we have *stado ovec* 'flock of sheep' where *ovca* 'sheep' is feminine, and *stado gusej* 'flock of geese' where *gus'* 'goose' is masculine. For number, we have *massa benzina* 'a mass of petrol' and *massa ljudej* 'a mass of people'. For case, we find that the head noun changes in case, but the dependent noun remains in the genitive, e.g. nominative *stado ovec*, dative *stadu ovec*, instrumental *stadom ovec*. Finally, the head noun does not change depending on the animacy of the dependent noun, cf. accusative *massu ljudej* 'mass of people' and *massu karandašej* 'mass of pencils'.

On the other hand, the head noun of a quantitative construction has a number of properties that are not shared by the adjective in the attributive construction, as follows: (e) the head noun can vary in number independently of the dependent noun, e.g. *stado ovec* 'flock of sheep', *stada ovec* 'flocks of sheep'; (f) the head noun in the quantitative construction can take an attribute agreeing with it, e.g. *xorošeje stado ovec* 'good flock of sheep', where *xorošeje* is neuter singular nominative, agreeing with *stado*, while *ovec* is genitive plural; (g) the noun dependent on the head noun is invariably in the genitive, and if countable in the genitive plural – contrast the attributive construction under point (c), where adjective and head noun must be in the same case.

In terms of their adherence to the above seven criteria, we find that we can divide Russian numerals into several classes. First, the numeral 'one' has all the properties of an adjective and none of those of a head noun: it

can even agree in number, with *pluralia tantum*, e.g. *odni* (PLURAL) *nožnicy* 'one (pair of) scissors'. At the other extreme, the numeral *million* 'million', and also all higher numerals, have all the properties of a noun and none of those of an adjective. Intermediate numbers have a varying number of adjectival and nominal properties, as illustrated in the table. In this table, A means that the numeral has the appropriate adjectival property, N that it has the appropriate substantival property; A/N means that either property can be used, A/(N) indicating that there is clear preference for adjectival behaviour; (A) means that the numeral has the adjectival property, but in a restricted form, in particular the numeral 'two' has only a two-way gender opposition, distinguishing feminine *dve* from masculine-neuter *dva*; (N) indicates a similar restriction on a substantival property, as with the plural of *sto* 'hundred', which has only a few restricted uses. In the table, note that 'four' behaves like 'three', and that non-compound numerals between 'five' and 'ninety' inclusive behave like 'five'.

ADJECTIVAL AND SUBSTANTIVAL PROPERTIES OF RUSSIAN NUMERALS

Property	*odin* '1'	*dva* '2'	*tri* '3'	*pjat'* '5'	*sto* '100'	*tysjača* '1000'	*million* '1,000,000'
(a)	A	N	N	N	N	N	N
(b)	A	N	N	N	N	N	N
(c)	A	(A)	N	N	N	N	N
(d)	A	A/(N)	A/(N)	N	N	N	N
(e)	A	A	A	A	(N)	N	N
(f)	A	A	A	A	A	N	N
(g)	A	A	A	A	A	A/N	N

If we now ask the question whether Russian numerals are adjectives or nouns, it becomes clear that there is no straightforward answer, except in the case of 'one' (adjective) and 'million' (noun): in particular, we cannot establish a cut-off point between adjectives and nouns, except arbitrarily, i.e. by deciding arbitrarily that we are going to take one, rather than another, of the seven criteria as definitive – and even then, some of the individual criteria are not definitive, as indicated by alternative entries separated by a slash or entries in parentheses. The situation is rather that we have clear prototypes, and a continuum separating those prototypes from one another, much as with colour terms, even though here we are clearly dealing with grammatical categories.

Actually, the continuum-like nature of the distinction between adjectival and substantival properties finds an even stronger manifestation in

Russian numerals if we also take into account statistical preferences where alternatives are possible. For instance, after the numerals 'two', 'three', and 'four', an adjective may be in either the nominative plural (as would be expected if these numerals were adjectives) or the genitive plural (as would be expected if these numerals were nouns). If one counts the occurrences of either possibility in text, it turns out that the preference for the adjectival type is greatest with 'two' and lowest with 'four', i.e. even as between adjacent numerals one can establish that the lower is more adjective-like than the higher.

In conclusion, definitions based on prototypes must be allowed as a possibility.

5.3 ERGATIVITY

In section 5.1, we posed a general problem for the syntactic analysis of any sentence, namely: what is the subject of the sentence? In view of the discussion of section 5.2, we can slightly reformulate that question. Implicit in the original question was that the question would have a clear-cut, discrete answer, i.e. a given noun phrase either would or would not be a subject. However, in terms of our characterization of subject as the intersection of agent and topic, and given that agent and topic are logically independent notions and need not coincide in a given sentence, it is clear that the answer to our question may well be less than clear-cut: it may be the case that a given noun phrase has certain subject properties, but not all, i.e. instead of simply saying that a noun is or is not a subject we will characterize it as being a subject to a certain degree. Similarly, it is possible that subject properties in a sentence will be distributed among several noun phrases, or at least between two, rather than all characterizing a single noun phrase. In many instances, then, it is as pointless to expect a clear-cut answer to the question 'what is the subject of this sentence?' as it is to expect a clear-cut answer to the question 'is Russian *pjat'* 'five' a noun or an adjective?' In the present section we will examine implications of this further, with particular regard to ergativity.

In section 5.1, we also posed the more specific question of identifying the subject of the ergative construction. In order to discuss this construction adequately, especially in terms of its similarities to and differences from the nominative–accusative construction, it is necessary to have a set of terms that is neutral between the two systems. The following is the set that we propose: The single argument of an intransitive predicate we will symbolize as S; this is clearly mnemonic for subject, and in general there is little or no controversy concerning the subject status in most intransitive

(single-argument) constructions across languages, so the mnemonically suitable symbol is also suitable in terms of its content. In the transitive construction, there are two arguments, and in order to avoid circularity we shall label neither of these with the symbol S. In the prototypical transitive situation, the participants are an agent and a patient, and this remains constant irrespective of the morphological or syntactic behaviour of the sentence in any individual language. We may therefore, starting originally with transitive predicates describing actions, label the agent as A, and the patient as P, so that in the sentence *I hit you*, or in its translation into Chukchi, irrespective of the case marking of the various noun phrases *I* will be A and *you* will be P. The labels are again clearly mnemonic, for agent and patient, respectively. However, the advantage of having arbitrary labels A and P rather than actually using agent and patient is that we can continue to use the arbitrary symbols even when we pass beyond prototypical transitive situations (i.e. actions) to other constructions in the language that have similar morphology and syntax. In English, for instance, the transitive verb *see* behaves morphologically and syntactically just like the action transitive verb *hit*, so that although in *I saw you* the pronoun *I* is not, in terms of semantic role, an agent, we can still symbolize it as A. A and P are thus syntactic terms, whose prototypes are defined in semantic terms.

In discussing examples (1) and (2) introduced at the beginning of this chapter, then, we can say that in (1) Chukchi *ɣəm* and English *I* are Ss; in (2) Chukchi *ɣəmnan* and English *I* are As, while Chukchi *ɣət* and English *thee* are Ps. Moreover, in English one case is used to encode S and A – a case of this kind is called nominative; and another case is used to encode P – a case of this kind is called accusative. In Chukchi, one case is used to encode S and P – a case of this kind is called absolutive; another case is used to encode A – a case of this kind is called ergative. The discussion thus far has related essentially to morphology, and we return to ergative–absolutive and nominative–accusative case marking in chapter 6. It is now time to turn to syntactic properties of subjects.

From the remarks made hitherto about subjects in English, it should be clear that English treats S and A alike as subjects for syntactic purposes, certainly for those syntactic points discussed so far, and indeed for most others. We can illustrate this by means of examples using coordination, in particular coordination of clauses that share a noun phrase in common and where that noun phrase is omitted in the second conjunct. If we try and conjoin sentences (8), (9), and (10), taking a transitive clause and an intransitive clause, in that order, then it is clear that we can conjoin, with omission of the second occurrence of the coreferential noun phrase, only (8) and (9), and not (8) and (10):

The man hit the woman. (8)

The man came here. (9)

The woman came here. (10)

The man hit the woman and came here. (= (8) + (9)) (11)

Even though sentence (11) contains no overt S for the intransitive predicate *came here*, it is absolutely clear to the native speaker of English that the only possible interpretation for this sentence is that the man came here, even though the alternative interpretation 'the man hit the woman and the woman came here' would make perfect sense. In other words, in order to permit omission of a noun phrase from a second conjunct, English makes two requirements: (a) the semantic requirement that the two noun phrases be coreferential; (b) the syntactic requirement that the two noun phrases be either S or A. For syntactic purposes, English treats S and A alike, so subject in English means S or A.

We may contrast this situation with the situation that obtains in Dyirbal, with the translations of our three English sentences (8)–(10):

Balan dʸugumbil baŋgul yaṛaŋgu balgan. (12)
woman-ABSOLUTIVE man-ERGATIVE hit
'The man hit the woman.'

*Bayi yaṛa banin*ʸ*u.* (13)
man-ABSOLUTIVE came-here
'The man came here.'

*Balan dʸugumbil banin*ʸ*u.* (14)
woman-ABSOLUTIVE came-here
'The woman came here.'

*Balan dʸugumbil baŋgul yaṛaŋgu balgan, banin*ʸ*u.*
(=(12) + (14)) (15)
'The man hit the woman, and the woman came here.'

(In Dyirbal, nouns are usually accompanied by a classifier agreeing in class, including gender, and case with the noun; in the above examples, these are *balan*, *baŋgul*, and *bayi*.) Note in particular that (15) does not, and in Dyirbal cannot, have the meaning of English sentence (11): the two sentences in the two languages are crystal-clear in their interpretations to native speakers, though the interpretations happen to be different in the two languages. Dyirbal, like English, has two restrictions on coordination with omission of a noun phrase, but while the semantic restriction is as in

English (the two noun phrases must be coreferential), the syntactic restriction is different: in Dyirbal, the coreferential noun phrases must be S or P. Thus for syntactic purposes, Dyirbal treats S and P alike, as opposed to A, so that in Dyirbal the appropriate grammatical relation is one that groups S and P together, in other words subject in Dyirbal means 'S or P'.

Although it might seem that the syntactic difference follows the morphological difference between nominative–accusative morphology in English and ergative–absolutive morphology in Dyirbal (as can be seen by comparing examples (12)–(14)), it is important to emphasize that this is not the case. In English, the syntactic identification of S and A proceeds even with non-pronominal noun phrases, which do not have a morphological nominative–accusative distinction. In Dyirbal, personal pronouns of the first and second persons happen to have nominative–accusative case marking, a fact to which we return in chapter 6, but this does not affect the ergative–absolutive basis of the coordination construction:

*ŋad*ʸ*a ŋinuna balgan.* (16)
I-NOMINATIVE you-ACCUSATIVE hit
'I hit you'

*ŋad*ʸ*a banin*ʸ*u.* (17)
I-NOMINATIVE came-here
'I came here.'

*ŋinda banin*ʸ*u.* (18)
you-NOMINATIVE came-here
'You came here.'

*ŋad*ʸ*a ŋinuna balgan, banin*ʸ*u.* (19)
'I hit you, and you/*I came here.'

We should also note that not all languages pattern either like English or like Dyirbal. In Chukchi, for instance, in coordinate constructions the omitted S of an intransitive verb can be interpreted as coreferential with either the A or the P of the preceding verb:

*ətlə*ɣ *-e talayvənen ekək*
father ERGATIVE he-beat-him son-ABSOLUTIVE
ənkʔam ekvetɣʔi.
and he-left (20)
'The father beat the son, and the father/the son left.'

In Yidiny, as we saw in section 3.4, the preferred interpretation for an omitted S follows the morphology (coreferential with an absolutive or nominative noun phrase in the transitive clause), thus combining aspects of nominative–accusative and ergative–absolutive syntax, whereas Chukchi is completely neutral as between them, in this instance. One important point that the Yidiny material illustrates particularly clearly is that it is misleading to classify a language as being either ergative or not, rather one must ask: to what extent, and in what particular constructions is the language ergative, i.e. where does its syntax operate on a nominative–accusative basis and where does its syntax operate on an ergative–absolutive basis. In Yidiny, then, in the transitive construction, in some instances the A will have subject properties under coordination (example (44) of chapter 3), in other instances the P will have subject properties (example (43) of chapter 3), in yet other instances subject properties will be distributed between the two noun phrases (example (45) of chapter 3).

In common with many, but not all, languages, both English and Dyirbal have different syntactic means of encoding the same semantic roles, i.e. different voices. In English, for instance, we can take the transitive sentence (8), with *the man* as A and *the woman* as P, and rephrase it as a passive, an intransitive construction, in which *the woman* appears as S and *the man* as an oblique object (i.e. neither S, A, nor P):

The woman was hit by the man. (21)

Since *the woman* is S of (21), and also S of the intransitive sentence (10), it is possible to coordinate these two sentences together, omitting the coreferential S from the second conjunct, to give (22), which has exactly the same meaning as Dyirbal sentence (15):

The woman was hit by the man and came here. (22)

In Dyirbal, it is possible to take a transitive sentence like (12) (or, for that matter, (16)) and rephrase it so that 'the man' appears as an S, and 'the woman' as an oblique object, adding the suffix *-ŋay* to the verb. This kind of voice, whereby the A of the basic voice appears as an S, has in recent work on ergativity come to be called the antipassive voice:

Bayi yaṛa *bagun dᵸugumbilgu*
man-ABSOLUTIVE woman-DATIVE
balgalŋanᵸu.
hit-ANTIPASSIVE (23)
'The man hit the woman.'

In Dyirbal, it is then possible to conjoin (23) with the intransitive sentence (13), of which 'the man' is also S. For reasons that go beyond our concerns here, the only order in which this particular conjunction is possible is with the intransitive clause first:

Bayi yaṛa baninᵸu, bagul dᵸugumbilgu balgalŋanᵸu. (24)
'The man hit the woman and (he) came here.'

Thus we see that one of the functions of different voices in languages is to redistribute subject properties: in English, to enable what would otherwise be a P noun phrase to have subject properties (as an S); in Dyirbal, to enable what would otherwise be an A noun phrase to have subject properties (as an S).

We may close the discussion of this section by recapitulating the main points, and driving them home with one further example. While the assignment of subject is clear in most intransitive constructions, especially those that are literally one-place predicate constructions, in transitive constructions we may find subject properties assigned either to the A, in which case we have nominative–accusative syntax, or to the P, in which case we have ergative–absolutive syntax. Some languages show strong preference for one or the other – e.g. English is largely nominative–accusative, Dyirbal largely ergative–absolutive – while other languages are more mixed. In Chukchi, the infinitive construction works on the nominative–accusative system, with omission of the S or A of the infinitive, with the suffix *-(ə)k*:

Yəmnan *ɣət* *tite*
I-ERGATIVE you-ABSOLUTIVE sometime
məvinretɣət *ermetvi-k.* (25)
let-me-help-you to-grow-strong
'Let me help you to grow strong.'

Moryənan *ɣət* *mətrevinretɣət*
we-ERGATIVE you-ABSOLUTIVE we-will-help-you
rivl-ək *əməlʔo ɣeċeyot.* (26)
to-move all gathered-things-ABSOLUTIVE
'We will help you move all the gathered items.'

In (25), the S of 'grow strong' is omitted; in (26), the A of 'move' is omitted. In the negative participial construction, with the suffix *-lʔ* on the

verb in the participial form, the construction may be used to relativize either the S or the P of the participial clause, but not its A (unless the clause is antipassivized, as in (29)), with relativization then effectively of the S):

E *-tipʔeyŋe-kə* *-lʔ* *-in*
NEGATIVE sing NEGATIVE PARTICIPLE ABSOLUTIVE
ŋevəćqet *rayтəɣʔi.* (27)
woman-ABSOLUTIVE she-went-home
'The woman who was not singing went home.'

Iɣər a *-yoʔ* *-kə* *-lʔ* *-etə*
now NEGATIVE reach NEGATIVE PARTICIPLE ALLATIVE
enm -etə *mənəlqənmək.* (28)
hill ALLATIVE let-us-go
'Now let us go to the hill which (someone) didn't reach.'

En *-ayтat-kə* *-lʔ* *-a*
ANTIPASSIVE chase NEGATIVE PARTICIPLE ERGATIVE
qaa *-k* *ʔaaćek-a* *vinretərkəninet*
reindeer LOCATIVE youth ERGATIVE he-helps-them
ŋevəćqetti. (29)
women-ABSOLUTIVE
'The youth who does not chase the reindeer
is helping the women.'

(Note that in (29) the object of the antipassive verb stands in the locative case.)

5.4 SEMANTIC AND PRAGMATIC FACTORS

So far, we have not related splits between nominative–accusative and ergative–absolutive syntax to the distinction between those properties that are more properly correlated closely with agent, and those that are more closely correlated with topic, and it is to this discussion that we now proceed, although our discussion will necessarily involve only exemplification of a limited number of properties.

We may start off with subject properties that correlate more closely with agent properties. In many languages, in imperatives it is possible to omit reference to the addressee if that addressee is an A or an S, but not if it is a P; indeed, many languages have an even stricter requirement, namely that

the S or A of an imperative construction must be second person (addressee), i.e. they only have second person imperatives. This can be illustrated for English by the examples *come here!* (i.e. *you come here!*) and *hit the man!* (i.e. *you hit the man!*), where it is possible to omit the addressee pronoun, in contrast to *let/may the man hit you!*, where it is not possible to do so. Interestingly enough, in Dyirbal, precisely the same constraint holds: despite the widespread prevalence in this language of syntactic constructions where S is identified with P, in imperative addressee deletion S is identified with A, as in English:

(ŋinda) *bani.* (30)
you-NOMINATIVE come-here-IMPERATIVE
'Come here!'

(ŋinda) *bayi yaŋa* *balga.* (31)
you-NOMINATIVE man-ABSOLUTIVE hit-IMPERATIVE
'Hit the man!'

The motivation for this distribution is not hard to find. For an instruction to be felicitous, the person to whom the instruction is addressed must have control over the resultant situation. In general, S and, especially, A are the participants who have most control over the situation, whereas P rarely has much control, so that it is more natural that the recipients of instructions should be encoded linguistically as an S or an A than as a P. Imperative addressee deletion simply provides a more compact means of expression for the more expected situation, i.e. addressees can be deleted when they are the more agentive S or A, but not when they are the less agentive P. This is thus a clear instance of a subject property that correlates with an agent property. Note that we are not saying that subject and agent are identical with respect to this property, or that the syntactic rule can be stated in terms of agents rather than in terms of subjects. For English, this is clearly untrue, since one can form passive imperatives where the addressee is not an agent but can be deleted, or where the agent is addressee but cannot be deleted (although the resultant sentences are very unnatural):

Be amazed by the world's greatest lion-tamer! (32)

Let/may this problem be solved by you! (33)

What we are claiming is that this subject property has a high correlation with an agent property, and therefore the S/A identification is more natural, even in a language like Dyirbal where the S/A identification otherwise plays little or no role in the language.

Moreover, we are not claiming that a language will necessarily have S/A identification for a subject property that correlates highly with an agent property, only that there will be a strong tendency for this to be the case (i.e. a universal tendency rather than an absolute universal). In Dyirbal, for instance, one might expect the same nominative–accusative syntax to carry over to indirect commands, deleting the S or A of the indirect command if coreferential with the recipient of the command. In fact, however, the A of an indirect command cannot be deleted in that form, rather the antipassive must be used, presenting that noun phrase as an S, which can then be deleted by the general rule allowing deletion of either an S or a P:

ŋana yabu gigan ŋumagu
we-NOMINATIVE mother-ABSOLUTIVE told father-DATIVE
buṛalŋaygu. (34)
see-ANTIPASSIVE-INFINITIVE
'We told mother to watch father.'

(Note that the dative is one of the possible cases for the patient in the antipassive construction.) If the unmarked voice is used for a transitive verb in the infinitive, then only a coreferential P may be omitted, as in (35):

ŋaḍ'a bayi yaṛa gigan
I-NOMINATIVE man-ABSOLUTIVE told
gubiŋgu mawali. (35)
doctor-ERGATIVE examine-INFINITIVE
'I told the man to be examined by the doctor.'

The example of imperative addressee deletion involved a natural identification of S and A, i.e. natural nominative–accusative syntax. We may now turn to an example of natural ergative–absolutive syntax. In Nivkh, there is a resultative construction, i.e. a construction referring to a state that has come about as the result of a previous event, using the suffix *-ɣəta*. With intransitive verbs, this involves simply the addition of the suffix to the verb:

Anaq yo -d'. (36)
iron rust
'The iron rusted.'

Anaq yo -ɣəta -d'. (37)
iron rust RESULTATIVE
'The iron has rusted.'

(The verb-final suffix *-d'* is an indicator of finiteness.) If, however, we take a transitive verb, then a number of changes take place relative to the non-resultative form:

Umgu t'us t^{h}a -d'. (38)
woman meat roast
'The woman roasted the meat.'

T'us ř̆a -ɣəta -d'. (39)
meat roast RESULTATIVE
'The meat has been roasted.'

First, for the majority of transitive verbs in most circumstances, the A of the transitive verb must be omitted in the resultative construction. Secondly, the P of the transitive verb has the property that it conditions consonant-initial alternation in the verb (cf. the initial *t^h-* of (38)), and the absence of such alternation in the resultative verb suggests that this noun phrase is no longer P. Whatever the precise details of the syntactic analysis, we can say that the resultative verb has a single argument, and that this argument corresponds to the S of a non-resultative intransitive verb, but to the P of a non-resultative transitive verb. In other words, S and P behave alike, as opposed to A.

The explanation this time is to be sought in the pragmatic structure of resultative constructions. Any such construction attributes a change of state to a certain entity. With intransitive predicates, the change of state is necessarily attributed to the S: in sentence (37), it is the iron that has undergone a change of state. With transitive predicates, although it is in principle possible for the change in state to characterize the A, as in *John has climbed the mountain*, it is more usual, especially with the prototypical transitive predicates describing an action involving a change of state, for the change of state to be attributed to the P. If we say *the woman has roasted the meat*, then we are necessarily talking about a change of state in the meat, and whether or not there is any change of state in the woman is simply left open. What Nivkh does is to grammaticalize this natural topicalization of S or P in the resultative construction, by allowing only S or P to be expressed.

Again, we are not claiming that a language must make this identification in the syntax of resultative constructions. English, for instance, does not, so that *the woman has roasted the meat* is perfectly acceptable as the resultative of *the woman roasted the meat*. We are claiming, however, that languages will tend to show a bias towards ergative–absolutive syntax in resultative constructions.

In many constructions, unlike imperatives and resultatives, there seems, a priori, to be no expected bias towards identifying S with either of A or P, for instance with coordination omission of noun phrases, and it is in these constructions that we find most variation across languages: with coordination, for instance, English has nominative–accusative syntax, Dyirbal has ergative–absolutive syntax, Yidiny has both, and Chukchi has neither. However, our present understanding of the cross-language distribution in such cases suggests that nominative–accusative syntax is in fact more widespread than ergative–absolutive syntax, and we might ask why this is so. Moreover, if we take a piece of natural nominative–accusative syntax like imperative addressee deletion, there are few or no languages that go against it by having ergative–absolutive syntax. However, if we take a piece of natural ergative–absolutive syntax, like resultative constructions, then we do find a wide range of languages that go against the natural syntax by having nominative–accusative syntax. In other words, there seems to be a general bias in language, interacting with naturalness of identification of S with A or P, towards nominative–accusative syntax. This general bias, in turn, has an explanation: as we shall see in a slightly different context in chapter 9, humans have a strong tendency to select more agentive entities as topics of discussion, which means that there is a natural correlation between agent and topic: other things being equal, one would expect agent and topic to coincide. The notion of subject then simply reflects the grammaticalization of this expected coincidence, and explains why so many languages do have a grammatical relation of subject definable in its core as the intersection of agent and topic, whereas few languages similarly define grammatical relations reflecting the intersection of, say, patient and topic.

While preference for equating agent and topic does seem by far the most prevalent identification across languages, there are some languages that do not show this particular identification. In Dyirbal, for instance, subject properties that are not agent-bound, and even some of those that are (cf. indirect commands), adhere to the P rather than to the A. In Dyirbal, then, it seems that agentivity is virtually irrelevant to the establishment of subjecthood, preference being given to P, i.e. P is more natural as a topic than A. In a number of Austronesian languages, especially in Philippine languages, a similar, though somewhat less extreme, situation seems to obtain, with some syntactic processes being conditioned by the agentivity (semantic role) of the noun phrases involved, and other syntactic processes – including most of those where no bias would be expected a priori – being controlled by the topic, with preference for a patient rather than an agent to be topic. The following examples are from Tagalog.

If we take the situation of someone borrowing money from a bank, then in Tagalog if the P is definite, it must be topic, irrespective of the definiteness of the other noun phrases, as in (41). Only otherwise is it possible for the A to be topic, as in (40):

Humiram siya ng pera sa bangko. (40)
borrowed-ACTIVE he-TOPIC P money DATIVE bank
'He borrowed money from the bank.'

Hiniram niya ang pera sa bangko. (41)
borrowed-PASSIVE he-A TOPIC money DATIVE bank
'He borrowed the money from the bank.'

If we embed this construction under a verb meaning 'to hesitate', then this is a construction which, a priori, favours S/A identification – one can only hesitate about something under one's own control – and here Tagalog allows deletion of the A, irrespective of whether it is topic or not:

Nagatubili siya -ng humiram ng pera sa bangko. (42)
hesitated-ACTIVE he-TOPIC
'He hesitated to borrow money from the bank.'

Nagatubili siya-ng hiramin ang pera sa bangko. (43)
'He hesitated to borrow the money from the bank.'

(In the last two examples, the suffix *-ng* is a clause-linker.)

If, however, we take a construction that is neutral as between identification of S with A or P, then Tagalog treats the topic as subject. For instance, in Tagalog relative clauses, the noun relativized can only be topic of the relative clause. Compare the following sentences:

Bumili ang babae ng baro. (44)
bought-ACTIVE TOPIC woman P dress
'The woman bought a dress.'

Binili ng babae ang baro. (45)
bought-PASSIVE A woman TOPIC dress
'A/the woman bought the dress.'

If we want to relativize 'woman', then the active construction of (44) must be used, but if we want to relativize 'the dress', then the passive construction of (45) must be used; no other alternatives are possible:

Iyon ang babae-ng bumili ng baro. (46)
that TOPIC woman bought-ACTIVE P dress
'That is the woman who bought a dress.'

Iyon ang baro-ng binili ng babae. (47)
'That is the dress that the woman bought.'

To conclude this chapter, we note that treating subject as a diffuse, rather than a discrete, notion, while perhaps seeming at first to weaken the notion of subject, does in fact provide us with a powerful tool which, in conjunction with independently established correlations with agent and topic properties, enables us to describe in a unified way, with a large measure of explanation, disparate phenomena across a wide range of languages.

NOTES AND REFERENCES

The idea of defining prototypical subject as a multi-factor concept is developed initially by Keenan (1976b), although I do not use his classification of properties here. The strongest criticism of this approach comes from Johnson (1977a), but unfortunately Johnson begs the question by assuming that a definition must be in terms of necessary and sufficient conditions. My discussion of the continuum ('squish') from adjective to noun in Russian numerals is based closely on Corbett (1978).

The discussion of ergativity in section 5.3 is based on Comrie (1978b). Very similar ideas, though with certain differences in terminology, emphasis, and concept, are given independently by Dixon (1979); note in particular that Dixon uses O for my P, uses subject for a natural grouping of S and A, and uses pivot for a grouping of S with A or S with P in a particular language. The Dyirbal examples derive originally from Dixon (1972), though they reappear in almost every subsequent discussion of ergativity; example (35) was expressly collected for me by R. M. W. Dixon, who thereby earns a particular debt of gratitude, the more so since this example requires slight modification of pages 128–9 of Dixon (1979). Numerous studies on ergativity are gathered in Plank (1979); the Chukchi examples are from the contributions to this volume by Comrie (1979c, 226, 227, 229) and Nedjalkov (1979, 242).

Splitting subject properties between agent (role) and topic (reference) properties is developed, especially for Philippine languages, by Schachter (1976, 1977); the Tagalog examples are taken from the second of these. The Nivkh examples are from Nedjalkov *et al.* (1974). Factors controlling the distribution of nominative–accusative and ergative–absolutive syntax are discussed by Moravscik (1978b). The discussion of imperative addressee deletion is based on Dixon (1979, 112–14), that of resultative constructions on Comrie (1981).

6

CASE MARKING

6.1 THE DISCRIMINATORY FUNCTION OF CASES

In this chapter, we are going to look at one way in which consideration of data from a wide range of languages has enabled us to gain important new insights into a general linguistic phenomenon, insights that would probably not have been obtained solely by the investigation of a single language, and certainly not from the detailed, abstract analysis of English. If one looks at the accounts given of the uses of cases in traditional, and many non-traditional, grammars, there is usually the assumption – in many instances, justified – that the use of a given morphological case will correlate highly either with a given semantic role, or with a given grammatical relation. Thus the locative case is said to be the case for expressing location, the ablative for expressing motion away from, and so on; the nominative is described as being the case for the subject, the accusative for the direct object (or, in frameworks that eschew the distinction between semantic and syntactic cases, nominative correlates with agent and accusative with patient). In addition to case marking systems based on semantic and/or syntactic criteria, recent linguistic research has also uncovered languages where pragmatic criteria are important in assigning case, as in Japanese and Tagalog, for instance.

In addition, however, to languages where some or all of the cases can be accounted for in this way, there remains a set of recalcitrant data, where on the basis of semantic roles or grammatical relations or pragmatic roles there remain some cases that do not correlate directly with any syntactic or semantic or pragmatic role, but rather seem to be used for a given role, but only in certain, limited circumstances. The aim of this chapter is to investigate some of these examples, in particular examples concerned with subjects and direct objects (or, more accurately, with S, A, and P). The reason why this discussion fits well into our general discussion of universals and

typology is that the kinds of non-correspondence that we shall be looking at are found to recur in a wide variety of languages from different genetic and areal groupings, i.e. we are dealing with a significant phenomenon from the viewpoint of language universals. Moreover, not only can we establish a general pattern of similar distribution across languages, we can actually go a long way towards finding an explanation for this cross-language similarity.

We shall begin our discussion by considering the nominative–accusative and ergative–absolutive case marking systems, already introduced in passing in chapter 5. If we take S, A, and P as our primitives, and assume for the moment that we are restricting ourselves to languages that treat each of these three relations homogeneously, i.e. do not have different cases for different types of S, etc., then it is clear that there are not just two logically possible kinds of case marking system, but five. The nominative–accusative system groups S and A (nominative) together against P (accusative). The ergative–absolutive system groups S and P (absolutive) together against A (ergative). Both of these systems are widespread across the languages of the world. The neutral system would have the same form for all three primitives, but since this is tantamount to lack of case marking for these relations, it is not directly relevant to our considerations: as a system, it is, of course, widespread in the languages of the world, but most languages with this system have other means, such as verb agreement or word order, to indicate which noun phrase is A and which is P in the transitive construction. The fourth possible type, tripartite, would have distinct cases for each of the three primitives. The fifth type would group A and P together as against S.

The tripartite system is found, but is very rare. In a number of languages, as we shall see in more detail below, it is found with a subset of the noun phrases in a language, namely where nominative–accusative and ergative–absolutive systems co-existing in a language intersect. But there is only one language for which it is reliably reported that this tripartite system exists for all noun phrases in the language, namely Wanggumara. Thus we can say with confidence that this system is very rare across the world's languages. The last type, with A/P–S alignment, seems to be equally rare: the only reliable attestations known to us are for certain classes of noun phrases in certain Iranian languages, where it represents an intermediate diachronic stage in the breakdown of an earlier ergative–absolutive case marking system in the direction of a nominative–accusative system. The question arises immediately why, of four logically possible case marking systems, two should account for almost all the languages of the world that have a case marking system that consistently distinguishes among S, A, and P. If we compare the noun phrase arguments of intransitive and transitive constructions, as in (1)–(2) (irrespective of word order), then a possible motivation for this distribution emerges:

$$S \quad V_{intransitive} \tag{1}$$

$$A \quad P \quad V_{transitive} \tag{2}$$

In the intransitive construction, there is only a single argument, so there is no need, from a functional viewpoint, to mark this noun phrase in any way to distinguish it from other noun phrases. In the transitive construction, on the other hand, there are two noun phrases, and unless there is some other way (such as word order) of distinguishing between them, ambiguity will result unless case marking is used. Since it is never necessary, in this sense, to distinguish morphologically between S and A or S and P (they never cooccur in the same construction), the case used for S can be used for one of the two arguments of the transitive construction. The nominative–accusative system simply chooses to identify S with A, and have a separate marker for P; while the ergative–absolutive system chooses to treat S the same as P, with a separate marker for A. The tripartite system is unnecessarily explicit, since in addition to distinguishing A from P, it also distinguishes each of these from S, even though S never cooccurs with either of the other two. The A/P – S system is, from a functional viewpoint, singularly inefficient, failing to make the most useful distinction (between A and P), and making a useless distinction (between A and S, likewise between P and S). Whatever may be the value of functional explanations in general in linguistics and language universals in particular, here we do have a good example where the predictions of the functional approach appear to fit in very well with the observed distribution of case marking systems across the languages of the world.

In fact, the functional approach makes a further prediction that is borne out by actual distribution. In a case system where one of the two cases used for indicating these three primitives is formally less marked than the other, for instance where one of the forms is simply the stem of the noun in question whereas the other has some overt affix, it is nearly always the case that the formally unmarked item is used to indicate S, whence also A in the nominative–accusative system and P in the ergative–absolutive system. This is Greenberg's universal number 38: 'where there is a case system, the only case which ever has only zero allomorphs is the one which includes among its meanings that of the subject of the intransitive verb', although a very few counterexamples to this generalization have since been uncovered, all with a nominative case more marked than the accusative, e.g. in Yuman languages, where the nominative takes the suffix -č and the accusative has no suffix. If, however, we restrict ourselves to the more

general pattern, then we can see that in the nominative–accusative system, a special marker is added to P to distinguish it from A, which like S is unmarked. In the ergative–absolutive system, a special marker is added to A to distinguish it from P, which like S is unmarked. The functional explanation of these two case marking systems may also explain why there is so often a discrepancy between the case marking system and the syntactic orientation of the language in question, as discussed in chapter 5: the cases do not relate directly to grammatical relations, but rather directly to distinguishing between A and P.

We would emphasize one point before proceeding further with the functional model of case marking and its implications. We are not claiming that the sole function of case marking is discriminatory in this sense, since there is a whole host of instances where the function of a given case can be correlated with semantic parameters. What we are claiming is that there do exist many instances where this functional approach is necessary in order to guarantee a full understanding of the role of case marking.

6.2 NATURAL INFORMATION FLOW IN THE TRANSITIVE CONSTRUCTION

From section 6.1, it emerges that the discriminatory function of case marking will show itself most clearly in the transitive construction, where there is a need to distinguish between A and P, rather than in the intransitive construction, where S alone occurs. Where one finds different cases used for different occurrences of S in a language, the conditioning factor is usually semantic (to the extent that it is not lexically idiosyncratic): for instance in Bats, as discussed in chapter 3 (sentences (1)–(2)), the distinction between the ergative and absolutive cases for intransitive subject is dependent on the degree of control exercised by the S over the situation described. There are also instances where differential case marking on A and/or P can be readily handled in semantic terms without appeal to functional factors. For instance, in Finnish the P stands in the partitive case if only partially affected by the action (e.g. if only some of an entity is affected), but in a non-partitive case if totally affected:

Hän otti rahaa (PARTITIVE). (3)
'He took some money.'

Hän otti rahan (ACCUSATIVE). (4)
'He took the money.'

In this section, however, we will be concerned with formal case distinctions that do not correlate this closely with a combination of semantic or syntactic factors, in particular trying to account for the following facts: a large number of languages have special cases for animate and/or definite Ps, distinct from the cases used for other Ps, and also not used elsewhere as markers of definiteness; conversely, many languages have a special case used only for As of low animacy, and not otherwise used as indicators of either A or low animacy.

Before proceeding to the data here, we will outline the explanation, following on from the discussion of the preceding section, that we will be appealing to, as this will make the citation of the individual pieces of data more comprehensible. In the transitive construction, there is an information flow that involves two entities, the A and the P. Although in principle either of A and P can be either animate or definite, it has been noted that in actual discourse there is a strong tendency for the information flow from A to P to correlate with an information flow from more to less animate and from more to less definite. In other words, the most natural kind of transitive construction is one where the A is high in animacy and definiteness, and the P is lower in animacy and definiteness; and any deviation from this pattern leads to a more marked construction. This has implications for a functional approach to case marking: the construction which is more marked in terms of the direction of information flow should also be more marked formally, i.e. we would expect languages to have some special device to indicate that the A is low in animacy or definiteness or that the P is high in animacy or definiteness. This is precisely what we will try to document in the remainder of this section.

In the immediately preceding discussion, we have introduced the two terms animacy and definiteness. We will return to definiteness in more detail later on in this chapter, but for the moment we can work with the general definition of definiteness as the presupposition that the referent of a definite noun phrase is identifiable by the hearer; in terms of English structure, a definite noun phrase will either be a pronoun, a proper name, or a common noun introduced by the definite article or a demonstrative or other determiner. Animacy is a much more complex phenomenon, to which we return in chapter 9. For the moment, suffice it to say that a noun phrase is higher in animacy if it is to the left on a continuum some of whose main points are: first/second persons pronouns > other human noun phrases > animal noun phrases > inanimate noun phrases.

If a given transitive construction has to be marked to show that it does not correspond to the normal direction of flow of information, then there are (at least) three ways in which this marking could be made. First, one could mark the construction as a whole, say by marking the verb, to indi-

cate an unexpected constellation of A and P; we examine this possibility in section 6.2.1. Secondly, one of the noun phrases (or both of them) could be marked, say by having a special marker for unexpected As (those low in definiteness or animacy) and/or for unexpected Ps (those high in definiteness or animacy); such examples are discussed in section 6.2.2.

6.2.1 INVERSE FORMS

A number of languages have special verb forms to indicate whether the transitive action is initiated by an A higher in animacy than the P or lower in animacy than the P (with the third possibility, A and P equal in animacy, being treated arbitrarily as the one or the other). Perhaps the most famous instance of this in the linguistic literature is in the Algonquian languages, where one set of verb forms, the so-called direct forms, are used when the A is higher in animacy than the P, while the so-called inverse forms are used where the P is higher than the A. The actual animacy hierarchy of Algonquian languages takes the form: second person > first person > third person proximate > third person obviative. The distinction between two subtypes within third person, proximate and obviate, the former higher in animacy than the latter, guarantees that there will never, in fact, be a transitive construction where A and P are equal in animacy.

The examples below are from Fox, though the general principle holds for Algonquian languages as a whole. The suffix *-aa* in these examples indicates the direct form, while *-ek* indicates inverse form. The prefix *ne-* indicates first person: this illustrates another important property of the Algonquian verb forms, namely that the prefix invariably encodes the participant higher in animacy, irrespective of its grammatical role:

ne *-waapam-aa* *-wa.* (5)
1SINGULAR see DIRECT 3
'I see him.'

ne *-waapam-ek* *-wa.* (6)
1SINGULAR see INVERSE 3
'He sees me.'

6.2.2 DIFFERENTIAL MARKING OF A AND P

The most widespread indication of unnatural combinations of A and P across languages, however, is not by marking the verb, but rather by marking one or both of the noun phrase arguments. The following patterns in particular are found: (a) mark a P high in animacy, i.e. the accusative case is restricted to Ps that are high in animacy; (b) mark a P high in definiteness, i.e. the accusative case is restricted to definite Ps; (c) mark an A that is low in animacy, i.e. the ergative case is restricted to noun phrases that are low in animacy. Somewhat embarrassing is the absence of clear attestations of the fourth expected type, namely marking of an indefinite A; languages seem rather to avoid this particular construction by outlawing or discouraging transitive sentences with an indefinite A, either recasting them as passives or by using a presentative construction (like English *there is/are* ...). In English, although the sentences *a bus has just run John over* and *a bird is drinking the milk* are surely grammatical, more natural ways of expressing these pieces of information would be *John has just been run over by a bus* and *there's a bird drinking the milk*. In most languages that use the three methods outlined above for indicating less natural combinations of A and P, the case marking of A and P is determined independently, i.e. any A below a certain degree of animacy is marked ergative, irrespective of the P; conversely, any P above a certain degree of definiteness or animacy is marked accusative, irrespective of the A. This contrasts with the inverse verb forms discussed in section 6.2.1, where it is usually the relation of A to P that is important. Finally, before proceeding to detailed exemplification, we should note that there are some languages where the occurrence of the special ergative or accusative marker is conditioned not by any specific rigid cut-off point on the animacy or definiteness hierarchies, but rather by a more general condition of the kind: use the special marker only if there is likelihood of confusion between A and P; the assessment of likelihood of confusion is left to the speaker in the particular context. Hua is an example of a language of this type.

For the relevance of animacy, particularly clear data are provided by Australian languages, almost all of which have split case marking determined by the animacy hierarchy. As would be expected from our discussion above, a special accusative case is often restricted to noun phrases towards the top of the animacy hierarchy: thus in Dyirbal it is found only with first and second person pronouns; in Arabana only with human noun phrases; and in Thargari only with animate noun phrases. Conversely, the special ergative case is found only towards the bottom of the hierarchy, though usually, in fact, in these languages extending quite high up the hierarchy: thus most Australian languages have a separate ergative case for all non-pronominal noun phrases (e.g. Dyirbal), sometimes extending further up the hierarchy into the pronouns. Since the determination of the case of A and P is independent, it sometimes happens that accusative and ergative case marking meet neatly in the middle of the hierarchy without any overlap or gap, but quite frequently there is overlap in the middle of the hierarchy, which means that some noun phrases have a tripartite case

marking system; and it sometimes happens that there is a gap in the middle of the hierarchy, some noun phrases having the neutral case marking system. Thus Ritharngu, for instance, has a nominative–accusative case marking system for pronouns; the tripartite system for humans and intelligent animals; and ergative–absolutive case marking for other nouns, i.e. for non-intelligent animals and inanimates. In some languages, the middle ground in the hierarchy may be shared by both the tripartite and neutral case marking systems, as was discussed in section 3.4 for the Saibai dialect of Kalaw Lagaw Ya, which thus combines within one language nominative–accusative, ergative–absolutive, tripartite, and neutral case marking.

One result of the split case marking pattern is that a single sentence, in addition to having a nominative A and an accusative P, or an ergative A and an absolutive P, can also have one of the patterns: ergative A and accusative P; nominative A and absolutive P. These possibilities were often effectively discounted in earlier work on ergativity, with its rigid distinction between nominative and ergative constructions. The following illustrations are from Dyirbal:

Balan dʲugumbil *baŋgul yaṛaŋgu* *balgan.* (7)
woman-ABSOLUTIVE man-ERGATIVE hit
'The man hit the woman.'

ŋadʲa *ŋinuna* *balgan.* (8)
I-NOMINATIVE you-ACCUSATIVE hit
'I hit you.'

ŋayguna *baŋgul yaṛaŋgu* *balgan.* (9)
I-ACCUSATIVE man-ERGATIVE hit
'The man hit me.'

ŋadʲa *bayi yaṛa* *balgan.* (10)
I-NOMINATIVE man-ABSOLUTIVE hit
'I hit the man.'

Although the most spectacular evidence for the relevance of animacy in the A does seem to come from Australian languages, it is also found in other languages. For instance, in some North-East Caucasian languages (e.g. Lak), nouns have an ergative–absolutive case marking system, but personal pronouns have a neutral system. This is particularly interesting in that it goes against an otherwise largely valid generalization that pronouns tend to distinguish more categories than do nouns.

The restriction of accusative marking to nouns that are high in animacy is very widespread across the languages of the world, and we will limit ourselves to a few examples. Even English provides relevant data here, since it has a nominative–accusative distinction with (many) pronouns, e.g. *I – me*, whereas it does not have any comparable distinction for other noun phrases. A particularly clear set of instances is provided by the Slavonic languages, where animacy is one of the key parameters determining whether a noun phrase will have a separate accusative case or not. In Russian, for instance, masculine singular nouns of the declension Ia have a separate accusative case (with the ending *-a*) if animate, but not otherwise:

Ja videl mal'čik-a/begemot-a/dub/stol. (11)
'I saw the boy/hippopotamus/oak/table.'

In Russian, all animate nouns in the plural have a separate accusative case, while no inanimate nouns do. In Polish, only male human nouns have a special accusative case in the plural, instantiating a different cut-off point on the animacy hierarchy:

Widziałem chłopców/dziewczyny/psy/dęby/stoły. (12)
'I hit the boys/girls/dogs/oaks/tables.'

The forms of the last four nouns are identical with the nominative plural, whereas the nominative plural of 'boys' is *chłopcy*.

There are data from a wide range of languages for special marking of definite direct objects: again, a few examples will suffice. In Turkish, only definite direct objects take the special accusative case suffix *-ı* (or its vowel harmony variants), all other direct objects being in the same suffixless form as is used for subjects (A or S):

Hasan öküz-ü *aldı.* (13)
Hasan ox ACCUSATIVE bought
'Hasan bought the ox.'

Hasan bir öküz aldı. (14)
Hasan a ox bought
'Hasan bought an ox.'

(In Turkish, *Hasan öküz aldı* is also possible, although it leaves open how many oxen were bought, i.e. 'Hasan bought an ox or oxen'.) In Persian, the suffix *-rā* is used to indicate definite direct objects:

Hasan ketāb-rā did. (15)
Hasan book ACCUSATIVE saw
'Hasan saw the book.'

Hasan yek ketāb did. (16)
Hasan a book saw
'Hasan saw a book.'

(As in Turkish, Persian also allows *Hasan ketāb did* 'Hasan saw a book or books'.)

What is particularly interesting in this respect is that some languages, in determining whether or not a P is to take the special accusative form or not, use both parameters of animacy and definiteness. In Hindi, for instance, a human direct object will normally take the postposition *ko* whether or not it is definite; only occasionally, and with affective value, does one find indefinite human noun phrases without *ko* in P position. Non-human, especially inanimate, Ps, however, never take *ko* if they are indefinite, though they may, and usually do, take *ko* if they are definite:

Aurat bacce ko bulā rahī hai. (17)
woman child ACCUSATIVE calling PROGRESSIVE is
'The woman is calling the/a child.'

?Aurat baccā bulā rahī hai. (18)

(The oblique form *bacce*, of *baccā*, is automatic before a postposition.)

Un patrõ ko paṛhie. (19)
those letters ACCUSATIVE read-POLITE
'Please read those letters.'

Ye patr paṛhie. (20)
these letters read-POLITE
'Please read these letters.'

Patr likhie. (21)
letters write-POLITE
'Write letters please.'

Thus, in order to know whether to assign *ko* to a P in Hindi, one must weigh against one another its position on both animacy and definiteness hierarchies, and even then there is room in the middle for subjective judgement.

A somewhat similar situation is observed in Spanish, in connection with the use of *a* to mark certain direct objects. Normally, this preposition is only used for human Ps, but such Ps must moreover be high in definiteness: in particular, human Ps that are non-specific occur without the preposition:

El director busca el carro/al empleado/a un empleado/un empleado. (22)
'The manager is looking for the car/the clerk/a clerk/a clerk.'

In this example, the difference between *a un empleado* and *un empleado* in P position is that the former implies that there is some specific individual that the manager is seeking, whereas the second implies simply that he needs any clerk.

Although we have treated animacy and definiteness as if they were unproblematic categories in the brief preceding discussion, this is in fact far from the case. In chapter 9, we return to examining animacy in more detail, but to conclude the present chapter we will turn to some problems concerning definiteness. One problem when we compare categories across languages is that we should have some basis on which to identify the same category in different languages. Thus, if we say that definite direct objects go into the accusative case in both Turkish and Persian, then we should be able to justify using the same term definite in referring to both these languages, and also to English, where the category definiteness exists but does not condition case marking. Failure to ensure this cross-language comparability would mean that we are not doing language universals research, but are simply analysing each language as an independent unit – and, unlike those linguists who maintain that this is the only way to study languages, we would be doing so surreptitiously by pretending, through use of the same term, that our results are comparable across languages. We will show below that a problem of this kind seems to arise in connection with definiteness, but that a solution to this problem is in fact forthcoming, a solution which, moreover, actually strengthens the universal base of our discussion.

The problem is that certain Ps in Persian and Turkish stand in the accusative case even though they are clearly not definite. In Persian, for instance, if one wants to say 'give one of them to me', then although the noun phrase 'one of them' is clearly, by definition, indefinite, yet still Persian here requires the definite marker *-rā*:

Yeki az ānhā-rā be man bedehid. (23)
one of them ACCUSATIVE to me give

In sentences (14) and (16) we illustrated the absence of the accusative marker in Persian and Turkish with the indefinite article *yek* or *bir*. However, although the direct object introduced by the indefinite article is clearly indefinite, both languages allow the accusative suffix here, so that the full range of data is actually:

Hasan bir öküz aldı. (24)

Hasan bir öküz-ü aldı. (25)

Hasan yek ketāb did. (26)

Hasan yek ketāb-rā did. (27)

The existence of the second example in each language might seem to quash any possibility of identifying the concept called definite in these languages with that called definite in the discussion of English.

An indication of the route out of this dilemma is, however, indicated by our discussion of animacy. Animacy is clearly not a single dichotomy between animate and inanimate, but rather a continuum along which we can range entities according to their degree of animacy, so that for instance people are more animate than animals, and animals more animate than inanimate objects. In describing definiteness cross-linguistically, we can make use of a similar notion of continuum, i.e. a continuum of definiteness (or specificity). Definiteness in the highest degree means, as in English, that the speaker presupposes that the hearer can uniquely identify the entity being spoken of. In Persian example (23) we are clearly not dealing with definiteness in this extreme degree, rather what is at issue is that the referent of the noun phrase has been delimited by specifying a certain set, which can be identified (namely *ānhā* 'them'), and then indicating that the entity which is to be given, while not uniquely identifiable, must still be a member of this identifiable set. This can be described by the term definite superset, meaning that the identity of the entity is not determinable absolutely, but some headway can be made in identifying it because it must be a member of a delimited set.

Turkish example (25) and Persian example (27) represent a different realization of the notion degree of definiteness/specificity. Although both members of each pair of sentences in (24)–(27) are translated the same way into English, they are far from equivalent in the original languages. The versions with the accusative marking on the P noun phrases suggest that the reference of the noun phrase in question is important, relevant for the discourse as a whole. In other words, in a discourse that started with (25) or (27) we would expect the ox or the book to recur in the discourse. The versions without the accusative suffix, however, are quite neutral in this

respect, and could be used, for instance, in simply relating the various events that happened to Hasan, without any particular interest in the ox or the book. We can refer to this distinction as relevance of referent identification. The absence of the accusative suffix advises the hearer not to bother about identifying the referent, while presence of this suffix advises him that the referent of this noun phrase, though not yet determinable by the hearer, will be of relevance to the ensuing discourse. So all uses of the accusative case can be linked together in terms of a hierarchy of definiteness: at one extreme we have complete identifiability of the referent; further down the hierarchy we have partial identifiability (definite superset); and further down still we have indication that identification of the referent is relevant; at the bottom, identification of the referent is neither possible nor relevant. If we then compare accusative case marking in Persian and Turkish with definiteness (say, the occurrence of the definite article with common nouns) in English, then we see that the same parameter is involved throughout, only the cut-off points are different in the various languages.

6.3 SUMMARY

To conclude this chapter, we may note that case marking, which has so often been viewed as an area of language-specific idiosyncrasy, often lacking in generalization even internal to a single language, can be the subject of fruitful language universals, fruitful not only in the sense that they involve cross-language generalizations about case marking, but also because they point the way to more adequate analyses of other areas of language structure.

NOTES AND REFERENCES

The discussion of the five homogeneous systems for case marking of S, A, and P is taken from Comrie (1978b, 330–4). The Wanggumara data are discussed by Blake (1977, 11). The A/P – S system, considered unattested by Comrie (1978b), is documented by Payne (1979, 443) for Roshani. My information on Yuman languages is from Pamela Munro (University of California, Los Angeles).

The presentation in section 6.2 stems from some of the ideas contained in Comrie (1978b, 384–8), as modified by the similar results obtained independently by Silverstein (1976). The explanation has been modified slightly in the direction of ideas presented in DeLancey (forthcoming). In particular, as noted by DeLancey, and also by Hopper & Thompson