

Relationships within the clause

After introducing the ways in which languages can indicate grammatical relations within the clause (Section 6.1), this chapter outlines in detail the major systems: constituent order (6.2); case systems (6.3); and agreement and cross-referencing (6.4). Section 6.5 looks at grammatical relations cross-linguistically, and asks whether there are universals. Section 6.6 is a case study of languages with ‘free’ word order, based on Warlpiri.

6.1 INDICATING GRAMMATICAL RELATIONS IN THE CLAUSE

In this chapter we investigate the relationships between verbs and their noun phrase (NP) arguments within the clause. All languages have intransitive clauses – clauses with a verb and just one NP participant, such as *The dog growled* – and transitive clauses – clauses with a verb and two NP participants, such as *The dog licked my friend*. The NP participants that occur in these basic clause types are known as CORE ARGUMENTS, and this chapter examines the ways in which the world’s languages distinguish between core arguments.

There are three main ways in which a language may indicate the relationship between core NPs and the verbal predicate. First, each core NP may have a fixed position in the clause: such a system uses CONSTITUENT ORDER to indicate the relationship between NP participants and verb. In English, both subjects and objects have a fixed position, which is how we determine who killed who in a pair of sentences like *The snake killed the bird* and *The bird killed the snake*.

But core NPs don’t have a fixed position in all languages. Core NPs in Latin can appear quite easily in different positions; both sentences in (1) have the same meaning, although the order of the NPs is different in (1a) and (1b):

- | | | | |
|-----|----|--------------------------------------|---------|
| (1) | a. | Puer-um puella audi-t. | (Latin) |
| | | boy-ACC girl.NOM hear-PRES.3SG | |
| | | ‘The girl hears the boy.’ | |
| | b. | Puella puer-um audi-t. | |
| | | girl.NOM boy-ACC hear-PRES.3SG | |
| | | ‘The girl hears the boy.’ | |

This variation in constituent order is possible because in Latin, the form of the NPs themselves indicates what relationship they have with the verb: this is CASE MARKING.

The third way in which a language can indicate the relationship between NP participants and the predicate is by verb AGREEMENT or CROSS-REFERENCING. Latin and English have a limited amount of agreement: the *-t* suffix on *audit* and the *-s* suffix on *hears* both indicate a third person singular subject (in English, this occurs only in the present tense). But many languages have far more extensive systems to indicate the participants via marking on the verb itself, typically in the form of pronominal affixes. Look back at Section 4.3.3.2: the head-marking language Kambera is a typical CROSS-REFERENCING language, where free pronouns are only used for emphasis or disambiguation. So in (2), there are no independent free pronouns meaning ‘I’ and ‘you’, and it’s only the subject and object markers on the verb that determine who’s doing the asking.

- (2) Jàka **ku-karai-kai** tiang ... (Kambera)
if 1SG.SU-ask-2PL.OBJ later
'If I ask you (*plural*) later ...'

In (2), the bound pronominal affixes on the verb (shown in bold) are clearly crucial, whilst in Latin and English, verbal agreement markers don't have much of a function in distinguishing subject and object.

These three systems – order, case and agreement – are not mutually exclusive: most languages use some combination of systems, although it is common for one to predominate. Sections 6.2 through 6.4 examine each system in turn. Section 6.5 then looks at noun phrases in terms of their GRAMMATICAL RELATIONS, or GRAMMATICAL FUNCTIONS.

6.2 ORDER OF PHRASES WITHIN THE CLAUSE

6.2.1 Basic and marked orders

As we saw in Chapter 1, linguists often talk about the ‘word order’ of a particular language. In fact, this term refers not to single words but to the order of PHRASES, so a better term is CONSTITUENT ORDER. Here, I concentrate on the order of the three major constituents in a transitive clause: subject, object and verb. In many languages, including English, subjects are distinguished from objects by having a fixed position for each NP, in the ordinary, basic constituent order. Given the three constituents S, O and V, there are six logically possible variations, and indeed all six orders do occur as a basic constituent order amongst the languages of the world:

- (3) Mpša e-lomile ngwana. (Northern Sotho)
S V O
 dog SU-bit child
 'The dog bit a/the child.'
- (4) Müdür mektub-u imzala-dı. (Turkish)
S O V
 director.NOM letter-ACC sign-PAST
 'The director signed the letter.'
- (5) Tuigeann Bríd Gaeilge. (Irish)
V S O
 understands Bridget Irish
 'Bridget understands Irish.'
- (6) E kamatea te naeta te moa. (Gilbertese)
V O S
 3SG kill.3SG the snake the chicken
 'The chicken killed the snake.'
- (7) kaikuxi etapa-vâ toto, papa tomo (Apalai)
O V S
 jaguar kill-PAST 3PL father 3PL
 'They killed a jaguar, father's group.'
- (8) anana nota apa (Apurinā)
O S V
 pineapple I fetch
 'I fetch pineapple.'

In the examples above, the constituent orders shown are all reasonably uncontroversial: they represent the basic order, or one of the basic orders, found in each of the languages. So, for instance, we can say Northern Sotho is an SVO language, and Turkish is an SOV language. However, saying a language has a certain basic constituent order doesn't mean that it never has any other orders. For instance, English has a basic SVO order, as in *They adore syntax*, but we can also use an object-initial order, as in *Syntax, they adore*, to give particular emphasis to the direct object, in this case *syntax*. An order which is used like this to focus on a constituent is known as a MARKED (= non-basic) order.

In some languages, it is not easy to decide on a basic constituent order. First, two (or more) orders may be UNMARKED – equally neutral. For instance, some verb-initial languages such as Fijian, Tongan and Samoan (all Austronesian languages) are not clearly definable as either VSO or VOS: both orders are frequent. Languages which allow all of the six possible constituent orders are common; which order is actually chosen depends on pragmatic factors such as focus, and which constituent

is the topic of the sentence. Some languages with free constituent order do have one order which is clearly basic. So for example, the Slavonic languages Polish and Russian are SVO. The native American language Mohawk, on the other hand, has no single basic or dominant order; this is also a common pattern cross-linguistically. Some languages also have free or very unrestricted WORD ORDER in the most literal sense; we examine such languages in Section 6.6.

Second, some languages have a different order in root clauses and in subordinate clauses. For instance, a number of Germanic languages, including German and Dutch, have SOV order in embedded clauses but have unmarked SVO order in root clauses; see Section 3.2.4 for discussion of this phenomenon.

Third, it may not be possible to tell whether there's an unmarked word order because sentences don't typically contain independent subject and object NPs. This is typically the case in languages which are strongly HEAD-MARKING (see Section 4.3). The verb itself in such languages always has subject and object markers, as in (2), but in natural discourse there are very few clauses containing both a lexical subject and a lexical object NP (like *The bird killed the snake*), so we can't easily say what the order of S, O and V might be.

In instances like all these, the constituent order which is designated 'basic' often depends more on the theoretical allegiances of the linguist than on any properties of the language. The criteria linguists use to determine a basic constituent order include frequency, which means seeing how often each order occurs in a text, and neutrality, which means looking at sentences with no particular focus or emphasis. Native speakers also have strong intuitions about which order(s) are the most neutral, if any, and indeed whether or not word order changes make any difference to the meaning of a sentence.

6.2.2 Statistical patterns

The six basic constituent orders presented in Section 6.2.1 don't all have equal frequency. Statistically, we would expect to find the world's languages split evenly amongst the six possible orders. But in fact the basic orders SVO and SOV are by far the most frequent, between them covering around 80 per cent of the world's languages (roughly equally split). VSO is the only other major group, covering perhaps 9–12 per cent of languages, including Celtic (for example, Welsh and Irish), Semitic (for example, Biblical Hebrew and Classical Arabic) and Polynesian languages (such as Maori). Languages with the basic order VOS are much rarer, covering around 3 per cent of the world's total. As noted in Section 6.2.1, though, many verb-initial languages have both VSO and VOS as basic patterns. Both OSV and OVS were once thought not to exist as basic orders, and in particular the OSV order is extremely rare. But both are attested in the languages of the Amazon basin, as shown in (7) and (8). A certain amount of estimation is unavoidable in any figures given, not least because reliable information on basic constituent order is not always available.

Two major generalizations about constituent order in the world's languages emerge from the statistics. First, the vast majority of languages have subject-initial

order (SOV, SVO), and even if subjects are not absolutely clause-initial, they generally precede objects (SOV, SVO, VSO). In one large language sample (Tomlin 1986), 96 per cent of the languages have subjects before objects. Why might this be? Subjects appear to be more salient than objects, which may account for their initial position: subjects typically initiate the action expressed by the verbal predicate, are often agents of that action or at least in control of it, and are often the topic of the clause. On the other hand, objects are prototypically the theme or patient, the entity which is acted upon, and are less likely as topics.

Second, the majority of languages place V next to O (in either order): again, over 90 per cent of a typical sample of languages do this. Only two constituent orders lack the VO / OV grouping – the extremely rare OSV order and the much more frequent VSO order. In VSO languages, though, there are often alternative orders available which do place O and V together. For example, many VSO languages have an SVO alternative order (e.g. Arabic and Berber). And the Celtic languages, though generally considered to be VSO (like the Irish example in (5)), also have a very frequent auxiliary-SVO word order, as in (9). In this order, the subject precedes the other main elements in the clause and a transitive verb and its object are also grouped together into a VP. Note that the finite element in this clause is the initial auxiliary, and the lexical verb *péinteáil* ‘paint’ appears in its infinitival form lower down in the clause:

- (9) Bhí an fear [VP ag péinteáil cathaoir inné]. (Irish)
 Aux S V O
 was the man PROG paint.INFIN chair yesterday
 ‘The man was painting a chair yesterday.’

This grouping of O and V which predominates cross-linguistically gives support to the traditional two-way division of the clause into a subject and a predicate, which in turn contains the verb and its object (see 5.1.3).

Examination of large statistical samples of languages also reveals that the word order *within* constituents correlates with the order of the major constituents themselves (see, for example, Dryer 1991). In Chapter 4, I introduced the idea that languages fall into two basic groups, HEAD-INITIAL and HEAD-FINAL.

Head-initial order

- The verb *precedes* its objects and complement clauses.
- Adpositions are prepositions, giving [P NP] order in PPs.
- Complementizers (such as *that*, *if*, *whether*) precede the clause they select as complement.

Head-final order

- The verb *follows* its objects and complement clauses.

- Adpositions are postpositions, giving [NP P] order in PPs.
- Complementizers follow the clause they select as complement.

It turns out that OV languages (the largest group is SOV) are very generally head-final, whilst VO languages (SVO plus all verb-initial languages) are characteristically head-initial. For example, OV languages are far more likely to have postpositions than prepositions: in a typical sample (for instance, Dryer 1991) around 96 per cent of verb-final languages are postpositional. On the other hand, VO languages are typically prepositional: only around 14 per cent of SVO languages have postpositions, and only 9 per cent of verb-initial languages. Similarly, in VO languages, complementizers such as *if* and *that* virtually always precede their subordinate clause, as in English. Conversely, in around 70 per cent of OV languages, the complementizers follow the subordinate clause; see for example the Japanese examples in exercise 3 in Chapter 3.

To summarize Section 6.2, for some languages constituent order is the major way to distinguish the grammatical relations (subject, object etc.) in a sentence. We expect such languages to have a fairly rigid constituent order, as is true of English, for example. Other languages have much more freedom of constituent order. These are typically languages which have case marking and/or a well-developed system of verb agreement: both these features allow subjects to be distinguished from objects even if the NPs don't have a fixed position in the sentence. The following two sections look in detail at case marking and agreement, starting with an examination of case systems.

6.3 CASE SYSTEMS

6.3.1 Ways of dividing core arguments

In Chapter 4, I introduced the concept of a head and its dependents. We saw that the relationship between these elements need not be marked morphologically at all (for instance, it's not indicated in Chinese): such languages have neutral marking. But if it is indicated, this can be either by marking on the head (head-marking) or on the dependents (dependent-marking). In languages with CASE systems, the noun phrase dependents are marked to show their relationship with the head element in the phrase or clause. This section concentrates on the relationships between a head verb and its NP arguments; case marking shows, for example, which NP is the subject and which the object.

We've often used the terms 'subject' and 'object'. But do these terms apply equally well to all languages? In this section, we'll see that it is helpful to distinguish between different types of subjects, in order to describe case systems that occur outside the familiar European language families. I will divide the CORE ARGUMENTS of a verb as shown in Table 6.1, and use the abbreviations S, A and O to designate their grammatical relations (Dixon 1972, 1979, 1994).

Table 6.1

The core arguments

Subject of an intransitive verb	S
Subject of a transitive verb	A
Object of a transitive verb	O

For example:

(10) The snake(S) hissed.

(11) The chicken(A) bit the snake(O).

Though the label ‘S’ is associated with ‘subject’, it specifically refers just to the ‘single’ argument of an intransitive verb, as in (10): I recommend using ‘single’ as a mnemonic for S. The label ‘O’ is clearly ‘object’. And ‘A’ is for ‘agent’, which is the prototypical semantic role taken by the subjects of transitive verbs such as ‘bite’, ‘examine’ or ‘regurgitate’. All languages must have some way of distinguishing the transitive subject, A, from the object, O, so that we can tell who gets bitten in an example such as (11). In languages like English, fixed constituent order does this work. What, though, if the constituent order is free? One solution is to ensure that A has a different *form* from O: this is the role of case marking.

A logically possible way of distinguishing the three core arguments would, of course, be to have a different marking for each of them. Such a language would distinguish three different cases, one for S, one for A, one for O; an example is given as (26) below. However, this is actually an extremely unusual system, cross-linguistically. The reason for this is undoubtedly because a much more economical system is attainable, using just two case distinctions. Only A and O need to be marked differently. There are no clauses with both an S and an A: they can’t co-occur, because within any given clause the verb is either transitive or intransitive. Similarly, there are no clauses with both an S and an O: if the verb is intransitive, it just has an S, and not an O. So to achieve the most economical case system possible, there are two equally logical alternatives, both of which require just two case distinctions.

The first system marks S and A in the same way, and O differently. In other words, all subjects receive one case marking, and objects receive a different case. This is known as the *NOMINATIVE/ACCUSATIVE* pattern, and it occurs in most European languages (a notable exception is Basque). In Modern English, full noun phrases have no case marking, but we can see the relics of a previous nominative/accusative case system in the forms of the first and third person pronouns:

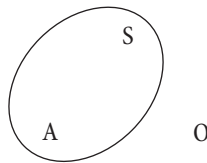
(12) We(S) left.
We(A) like her(O).

(13) She(S) left.
She(A) likes us(O).

We and *she* are NOMINATIVE forms, used for both S and A: in other words, all subjects have the same form. *Her* and *us* are ACCUSATIVE forms, used for O.

Figure 6.1

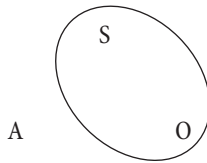
The nominative/accusative grouping



Because this grouping of S and A is so familiar from European languages, you may consider it entirely natural to case-mark all subjects in the same way. But remember that this is only one of the two equally economical ways of dividing the core arguments. The second system marks S and O in the same way, but marks A differently; this is known as the ERGATIVE/ABSOLUTIVE pattern:

Figure 6.2

The ergative/absolutive grouping



ERGATIVE is the case of A – the subject of transitive verbs. ABSOLUTIVE is the case of both S and O, the subject of intransitive verbs and the object of transitive verbs.

A summary of the two systems is shown in Table 6.2. You can see that both case systems only require two distinctions. One system groups S with A (since they never co-occur and so won't be confused with each other); this is typically known simply as the accusative pattern. The other system groups S with O (they, too, never co-occur, so also can't be confused with each other); this is typically known simply as the ergative pattern.

Table 6.2

The major case systems

Accusative system		
A	S	O
Nominative		Accusative

Ergative system		
A	S	O
Ergative	Absolutive	

In the following two sections I move on to an illustration of each of the main case systems in turn.

6.3.2 Nominative/accusative systems

I start with the most familiar system, NOMINATIVE/ACCUSATIVE (or just ACCUSATIVE). This system has an AS/O pattern: A and S are marked the same, O differently. Good examples are Latin, German, Japanese and Turkish, amongst many other languages. Subjects of both transitive and intransitive verbs are marked in the same way, with NOMINATIVE case. Objects of transitive verbs are marked with ACCUSATIVE case. This ‘alignment’ of NPs is sometimes indicated by using the notation $S = A \neq O$.

- (14) Puella veni-t. (Latin)
 girl.NOM come-PRES.3SG
 ‘The girl(S) comes.’
- (15) a. Puer-um puella audi-t.
 boy-ACC girl.NOM hear-PRES.3SG
 ‘The girl(A) hears the boy(O).’
 b. Puella puer-um audi-t.
 girl.NOM boy-ACC hear-PRES.3SG
 ‘The girl(A) hears the boy(O).’

Since the A and O arguments of the verb *audit*, ‘hears’, are in different cases, there is no problem determining which is which, despite the free constituent order illustrated in (15a) and (15b).

Case is generally considered to be a property of an entire *noun phrase*, rather than just the head noun itself. In some languages, case is indeed marked on the head noun via changes in its morphology (= changes in its form), as in the Latin examples. But elsewhere, for instance in German, case is typically not marked on the head noun, but is marked instead on the determiners and any adjectives in the noun phrase:

- (16) [Der gross-e Hund] knurrte. (German)
 the.NOM big-NOM dog growled
 ‘The big dog growled.’
- (17) [Der gross-e Hund] biss [den klein-en Mann].
 the.NOM big-NOM dog bit the.ACC small-ACC man
 ‘The big dog bit the small man.’

The (masculine) head nouns *Mann* ‘man’ and *Hund* ‘dog’ in (16) and (17) don’t undergo any morphological changes: they’re in their basic form. But we can tell who gets bitten in (17) from the case marking shown on other elements in the NPs, namely the determiners and the adjectives. For instance, *der* is the nominative form

of the definite article ('the') for masculine nouns, whilst *den* is its accusative form. The NP *den kleinen Mann* is thus shown as accusative, so it's the object, whilst *der grosse Hund* is nominative, so it's the subject.

6.3.3 Ergative/absolutive systems

The ERGATIVE/ABSOLUTIVE system (or just ERGATIVE) has an SO/A pattern: S and O are marked the same, and A is marked differently. Lezgian (a Daghestanian language spoken in the Caucasus) is a standard ergative language. The subject (A) of a transitive verb has ergative case, whilst the object (O) of a transitive verb and the subject (S) of an intransitive verb both have absolutive case. This ALIGNMENT of NPs is sometimes indicated by using the notation $S = O \neq A$. Compare in particular the forms of the first person singular pronouns ('I/me' in the English translations) in (18) through (20).

- (18) **Za** zi balk'an c'ud xipe-q^h ga-na. (Lezgian)
 I.ERG my horse.ABS ten sheep-for give-PAST
 'I(A) gave away my horse(O) in exchange for ten sheep.'

- (19) **Zun** ata-na.
 I.ABS come-PAST
 'I(S) came.'

- (20) Aburu **zun** ajib-da.
 they.ERG I.ABS shame-FUT
 'They(A) will shame me(O).'

In the English translations, the first person singular pronouns have the same form, *I*, both as an A and an S, whilst the O has a different form, *me*; this is the standard accusative case pattern. By contrast, in Lezgian the A form (*za*) differs from the S, and instead the S and O forms are identical (*zun*): this is the standard ergative case pattern. When the first person singular pronoun is an A – the subject of a transitive verb, as in (18) – it takes the ergative case, giving the form *za*. But when it's either an S (the subject of an intransitive verb) as in (19), or an O (an object) as in (20), it takes the absolutive case, giving *zun*.

Our second example comes from an ergative language spoken in Europe, namely Basque, which is a language isolate (= a language with no known relatives). Examples from the Lekeitio dialect are given in (21) through (23): compare the case marking of the word for 'man' in each example.

- (21) **Gixona-k** liburua erosi dau. (Basque)
 man-ERG book.ABS buy AUX.3SG
 'The man(A) has bought the book(O).'

(22) **Gixona** etorri da.
 man.ABS come AUX.3SG
 ‘The man(S) has come.’

(23) **Gixona** ikusi dot.
 man.ABS see AUX.1SG
 ‘I(A) have seen the man(O).’

The NP meaning ‘man’ has the ergative case suffix *-k* in (21), where it’s an A, i.e. the subject of a transitive verb. When this NP is an S or an O, as in (22) and (23), it takes the absolutive case, which has no actual suffix here but is instead the basic form of the noun.

If you understand the data, but are having difficulty remembering which NPs group together in the ergative/absolutive system, I recommend the mnemonic ‘Abso’, for ‘A-*but*-SO’ grouping. I hope this helps!

Ergativity is not found in the European language families (Romance, Germanic, Celtic, Greek, Albanian and so on – to which Basque is unrelated), and is also very rare in Africa. However, it is common in Australian languages, and also occurs widely in Tibeto-Burman languages, Mayan languages (Central America), and a number of Papuan languages (New Guinea), amongst others. In other words, ergative systems are not purely localized, but are spread around the world. Dixon (1994: 10) estimates that perhaps one quarter of the world’s languages can be described as ergative languages; clearly, then, the accusative system is far more common.

6.3.4 Split systems I

An important feature of all ergative languages is that they are never ergative in all aspects of their syntax and morphology, but instead have a combination of ergative and accusative properties. Often, a language doesn’t use just one case marking system consistently for all instances of A, S and O, but instead has ergative case marking for some constructions and accusative case marking for other constructions. The term in widespread use for such a system is **SPLIT ERGATIVE**. What this means is that the S argument may align with either A or O, depending on the grammatical context. In some circumstances, then, the alignment pattern is $S = A \neq O$ (an accusative alignment) and in other circumstances it is $S = O \neq A$ (an ergative alignment).

As an illustration of a split system, consider an Australian language, Dyirbal, which treats full noun phrases differently from pronouns in terms of case marking. Noun phrases are marked according to the ergative/absolutive system, as in (24). In Dyirbal, and very typically in other ergative languages, there is no actual inflection for the absolutive form; we saw this in the Basque data above. Here, the simple noun root is used for absolutive case, whilst the ergative is marked with a suffix, *-nggu*:

- (24) a. nguma banagan^yu (Dyirbal)
 father.ABS returned
 'Father(S) returned.'
- b. yabu banagan^yu
 mother.ABS returned
 'Mother(S) returned.'
- c. nguma yabu-nggu buran
 father.ABS mother-ERG saw
 'Mother(A) saw father(O).'

The word for 'father' has the same case, absolutive, when it's an S (24a) and when it's an O (24c). The word for 'mother' is an S in (24b), and so again has absolutive case, but it's an A (transitive subject) in (24c), so here it has the ergative case. However, pronouns in Dyirbal employ a different system, as you now have the opportunity to work out for yourself.



Before reading further, please examine the sentences in (25) and work out how the case-marking system for pronouns differs from that of full noun phrases.

- (25) a. ngana banagan^yu (Dyirbal)
 we.NOM returned
 'We(S) returned.'
- b. n^yurra banagan^yu
 you.NOM returned
 'You(S) returned.'
- c. n^yurra ngana-na buran
 you.NOM we-ACC saw
 'You(A) saw us(O).'



First and second person pronouns in Dyirbal have an accusative case marking system. So the S and the A pattern together: both are nominative, as in familiar European languages. The nominative form has no inflection, but just uses the bare root of the pronoun. The accusative form, the O, has an accusative suffix *-na*: compare the words for 'we' in (25a) and (25c).

Three major factors have been identified as responsible for triggering splits in the argument-marking system in ergative languages. First, properties of the NPs in the sentence (such as whether or not they are animate) can trigger a split. In the Dyirbal system, full NPs and third person pronouns employ the ergative system, (24), whilst other pronominals employ the accusative system, (25). An alternative found in some languages is that independent NPs exhibit an ergative alignment, whilst pronominal affixes employ the accusative system; see Section 6.4.4 on Warlpiri.

Note that 'If pronouns and nouns have different systems of case inflection, then the pronoun system will be accusative, and the noun system ergative, never the other way round' (Dixon 1994: 84).

Second, in some languages (e.g. Dakota) the meaning of an intransitive verb may trigger the split, according to how much control its S argument exerts over the verb's action. With volitional verbs like 'run' or 'swim', 'retire' or 'resign', the subject is semantically an agent, or at least has control over the verb's action or performs it voluntarily. This type of S argument will be marked in the same way as the A argument of a transitive verb, which is not surprising as both S and A are agent-like in such instances. But with non-volitional intransitive verbs like 'blush' or 'fall' or 'die', the S argument is not at all in control of the verb's actions; we blush or fall involuntarily, not deliberately. So this type of S argument will be marked in the same way as the O argument of a transitive verb, which is again not surprising as S is semantically much closer to a typical object in such instances: it *undergoes* the verb's action (blushing or falling etc.) rather than initiating it.

Third, the tense or aspect of the verb can trigger a split: ergative marking typically occurs with completed events, so is expected with past tense verbs or those with perfective aspect, whilst accusative marking occurs with present tense verbs and those with imperfective aspect. Hindi and various other Indo-Iranian languages illustrate this situation.

Dixon (1994) also reports that a fourth, much rarer, factor which may trigger splits in argument-marking is the status of the clause as either a main or an embedded clause. Only a few ergative languages are known to exhibit this type of split.

Finally, I noted in Section 6.3.1 that languages which use a different case for each of the core arguments, S, A and O, are very rare. Such a TRIPARTITE system does occur in another split-ergative Australian language, Pitta-Pitta, where the split is triggered (the technical term is ‘conditioned’) by the tense of the verb. In non-future tenses (but not in the future tense), NPs each have a different case, depending on whether they are S, A or O. In (26), look especially at the first person singular pronouns, the words for ‘I’ (in bold):

- (26) a. **nga-tu** kat^ʔu-na wat^ʔama-ka (Pitta-Pitta)
I-ERG clothes-ACC wash-PAST
'I(A) washed the clothes(O).'
- b. **angka-ya ngan^ʔɬa** kunti-ina.
sit-PRES I.NOM house-in
'I(S) am sitting in the house.'
- c. **tupu-lu ngan^ʔa** pat^ʔa-pat^ʔa-ya
caterpillar-ERG I.ACC bite-bite-PRES
'A caterpillar(A) is biting me(O).'

The pronoun for 'I' has ergative case in (26a), where it's an A; it has nominative case in (26b), where it's an S; and it has accusative case in (26c), where it's an O. These examples show that in Pitta-Pitta, the two main case marking systems – the ergative system and the accusative system – partially intersect.

Another kind of split ergative system is illustrated in Section 6.4.4 below.

6.3.5 Marked and unmarked forms

At this point we can discover why linguists often just use the terms ‘ergative’ or ‘accusative’ to describe the two systems: it is common for *just this one* member of each system to be the only NP that is overtly case marked, whilst the other member of each system is unmarked, i.e. has no special inflection for case at all. Instead we find the ordinary root of the noun or pronoun (the form with no inflections).

In an ergative system, if one form lacks overt marking it will be the absolutive NP, whilst the ergative NP has a special inflection. This is true of all the ergative systems illustrated so far, Lezgian, Basque and Dyirbal. Please confirm this by looking at the Dyirbal examples in (24): the absolutive forms are not inflected; the ergative form is.

In an accusative system, if one form lacks overt marking it will be the nominative NP, whilst the accusative NP has a special inflection. This is confirmed by (25): the nominative pronouns are not inflected, whilst the accusative one is.

In fact, we can make a generalization which works for both case systems: whichever case is used for the S argument (either absolutive or nominative), that will generally (with a few exceptions) be the NP that *lacks* any overt marking (Dixon 1994: 56f). Not only is the case used for S generally formally UNMARKED (= lacking special marking), as in the Dyirbal examples in (24) and (25), it’s also functionally UNMARKED. This means it’s more widespread in occurrence and more basic in terms of usage. For instance, the absolutive or nominative form is typically used as the citation form of a noun, generally the form given in a dictionary.

6.4 AGREEMENT AND CROSS-REFERENCING

6.4.1 What does verb agreement involve?

Case marking and verb agreement (also termed concord) are in fact two alternative (and sometimes overlapping) ways to represent the same information. Recall that the relationship between a head verb and its dependent NPs can be morphologically indicated either by DEPENDENT-MARKING (case) or HEAD-MARKING (agreement). As we saw in Chapter 4, it’s very common for a language to have both verbal agreement with a subject and also case marking on the core NPs: see example (1) from Latin. This is an instance of case and agreement overlapping.

In this section we’ll see that ‘an accusative system’ doesn’t necessarily imply that the language has nominative/accusative *case*; the same applies to ‘an ergative system’. The relationships between verb and core NPs may instead be shown by verb agreement, either following the accusative pattern or the ergative pattern. In other words, the verb agrees with certain of its dependents and not others.

Agreement, or cross-referencing, means that a head verb is formally marked to reflect various grammatical properties of its NP arguments; if you need to refresh your memory, look back at Sections 2.2.2.4 and 4.3.3.2. To take a simple example, a

verb might be marked for third person singular when its subject is a singular NP, and third person plural when its subject is a plural NP. (English has a verb agreement marker for third person singular subjects, -s, but only in the present tense: *S/he sings*). Cross-linguistically, the most common categories involved in agreement are PERSON, NUMBER, GENDER (= noun class) and CASE. We will see that verb agreement can follow an accusative or an ergative pattern even when there's no actual case marking on the NPs themselves.

Logically, the options are for a verb to agree (i) with none of its arguments, (ii) with some but not others or (iii) with all its arguments, and in fact all of these possibilities occur, as we will see below. There are, then, languages with no verb agreement whatever, for example Swedish, Japanese, Chinese, Maori and Malagasy. Example (27) illustrates this for Chinese:

- (27) a. Wo xihuan ta (Chinese)
I like he
'I like him.'
b. Ta xihuan wo
he like I
'He likes me.'

The verb has the same form, *xihuan*, no matter what the person and number of the subject pronoun. In fact, constituent order is the sole way of distinguishing the subject and object in these examples, since there's no case marking on the NPs either: the third person singular pronoun, for instance, is *ta* whether it's a subject or an object. Chinese illustrates a neutral alignment system.

We next turn to languages that do have verb agreement.

6.4.2 Nominative/accusative agreement systems

Within the Indo-European family, it is common for the verb to agree only with the subject, as for example in Italian, French, Spanish, German, Dutch and English. Subject-only agreement also occurs in Turkish and other Altaic languages, in Tamil and other Dravidian languages, and in Finnish and other Uralic languages. Examples (28) and (29) illustrate from French:

- (28) Nous avons vu ce film. (French)
we have.1PL seen this film
'We(A) have seen this film(O).'
- (29) a. Nous avons décidé.
we have.1PL decided
'We(S) have decided.'
b. Ils ont décidé.
they have.3PL decided
'They(S) have decided.'

These three examples show that there is subject/verb agreement in French, expressed on the auxiliary here rather than the lexical verb. So auxiliary *avons* has a first person plural inflection to agree with the 1PL subject pronoun *nous*, and a third person plural inflection *ont* to agree with the 3PL subject pronoun *ils* (29b). The subjects of transitive verbs (A) and the subjects of intransitive verbs (S) are both marked on the verb in the same way, whilst the verb does not agree with the object, *ce film*, in any way in (28). We can therefore say that French has an ACCUSATIVE agreement pattern, or alignment: A and S pattern together, as opposed to O. Again, this alignment can be indicated as $S = A \neq O$. French does not have case marking on NPs, however: as in English, only pronouns display the relics of an earlier case system.

The other possibility, also common cross-linguistically, is that the verb cross-references more than one of its arguments. So in Kambera, which also has an accusative alignment, the verb cross-references both the subject *and* the object: these markers are shown in bold type in (30). To help you see what refers to what, I've indicated both the independent subject NP and the bound subject marker on the verb with a subscript _{SU}. I also indicate both the independent object NP and the bound object marker on the verb with a subscript _{OBJ}:

- (30) [I Ama]_{SU} **na**_{SU}-kei-**ya**_{OBJ} [na rí muru]_{OBJ}. (Kambera)
 the father 3SG.SU-buy-3SG.OBJ the vegetable green
 'Father buys the green vegetables.'

In Kambera, the subject marker is a prefix (i.e. it precedes the verb stem), and the object marker is a suffix (i.e. it follows the verb stem). Example (30) has an overt subject and object, but if these are omitted the sentence is still perfectly grammatical, because the bound pronominals alone serve to indicate both a subject and an object. Such a sentence then simply has the (less specific) meaning 'He/she buys it'. Look back at example (2) above to see an instance of this kind. As we saw there, and in Chapter 4, head-marking languages (such as Kambera) often have whole sentences consisting of just the verb with its inflections. Free pronouns are generally not required, since the pronominal person and number affixes on the verb provide all the information about the verb's arguments: again, see (2).

In some languages, constituent order affects which agreement markers occur. So for example, in Northern Sotho, a Bantu language, the unmarked (= basic, usual) constituent order is SVO, as in (31) and (32):

- (31) Mpša e-lomilê ngwana. (Northern Sotho)
 dog SU-bit child
 'The dog bit a/the child.'
- (32) Di-mpša di-lomilê ngwana.
 PL-dog SU-bit child
 'The dogs bit a/the child.'

In (33) and (34) we have two variations on (31). These examples both have a marked constituent order, namely OSV in (33) and SOV in (34). And in these marked orders, we find both the subject marker *and* and an object marker, the prefix *mo-*. (This prefix agrees with the noun class of *ngwana*, ‘child’, which is Class 1, for human beings.)

- (33) Ngwana mpša e-mo-lomilè. (Northern Sotho)
O S V
child dog SU-OBJ-bit
'As for the child, the dog bit him/her.'
- (34) Mpša ngwana e-mo-lomilè.
S O V
dog child SU-OBJ-bit
'As for the dog, it bit the child.'

Before reading further, try to figure out why an object agreement marker is required in (33) and (34) but not in (31) or (32). Don't worry about the specifics of the noun classes or genders; this is not relevant to your answer.

First, consider (31) and (32): only one NP precedes the verb, so a Sotho speaker can assume that the order is the normal SVO order. The speaker can therefore tell that the first NP in the clause is the subject. Variations in this normal constituent order are used in Sotho to make a constituent the TOPIC of the sentence, with the TOPICALIZED NP appearing in initial position. The translations of (33) and (34) give the effect of this topicalization with the formula *As for the X*. In these examples, there are *two* NPs before the verb, either one of which might potentially be the subject. But since there's a subject marker *e-* which agrees with 'dog' and an object marker *mo-* which agrees with 'child', a Sotho speaker can sort out who's biting who. In this instance, the subject and the object are in different noun classes, so the sentence is completely unambiguous. Note that these subject and object agreement markers on the verb occur in a fixed order (in all languages, the order of elements *within* words is generally fixed), although either ordering of the independent object and subject NPs in (33) and (34) is grammatical.

If a language has object agreement, we can (with one or two exceptions) be sure that it will also have subject agreement: in other words, object agreement presupposes subject agreement. What about verbs which take more than two arguments, such as ditransitive verbs (see Chapter 4) like ‘give’ or ‘buy’? In some languages, a verb agrees with or cross-references more than two arguments, although this is not particularly common. In (35), from an Australian language called Biri, the verb cross-references three arguments, all of which are expressed as suffixes on the verb stem, and are shown in bold:

- (35) *nhula manhdha yaba-nha-la-nga-ngu* (Biri)
 he food give-FUT-3SG.SU-3SG.OBJ-1.DU.DATIVE
 ‘He will give food to us two.’

The verb stem in (35) is *yaba*, and this has a future tense marker, followed by three pronominal affixes, or person/number markers: *-la* marks the third person singular subject (and there’s also an independent third person subject pronoun *nhula*, ‘he’, here); *-nga* marks the third person singular object, agreeing with *manhdha* ‘food’; and *-ngu* is a marker for first person dual (‘us two’), and is also dative. Dative is a case often used to mark a recipient, which gives rise to the meaning of something being handed over to someone here.

In this section we have seen accusative systems of agreement: the verbs agreed with their subjects, or both with their subjects and objects.

6.4.3 Ergative/absolutive agreement systems

We turn now to systems with ergative alignment. When verb agreement follows the ergative pattern, it marks S (intransitive subjects) and O (all objects) in the same way and A (transitive subjects) differently. So we can say that $S = O \neq A$. ERGATIVE AGREEMENT MARKING occurs in a number of Caucasian languages, and also in Mayan languages (Mexico and Central America). Our examples are from the North-West Caucasian language Abaza. In (36) we have an intransitive verb, and in (37) a transitive verb. In all these examples, the data consist simply of a verb with bound pronominal affixes showing the person and number of the participant(s): I have indicated in bold the function of each morpheme:

- (36) a. *d-thád.* (Abaza)
 S-V
 3SG-go
 ‘He/she’s gone.’
 b. *h-thád.*
 S-V
 1PL-go
 ‘We’ve gone.’

- (37) a. h-l-bád.
O-A-V
 1PL-3SG.F-see
 'She saw us.'
- b. h-y-bád.
O-A-V
 1PL-3SG.M-see
 'He saw us.'
- c. d-h-bád.
O-A-V
 3SG-1PL-see
 'We saw him/her.'

All the person/number markers are prefixes on the verb in Abaza: note that they have a fixed order, S-V and O-A-V, so it is always clear who's doing what. The prefixes show the SO versus A pattern characteristic of ergativity. Throughout, any S and O markers which refer to the same person/number have the same form. First, let's look at third person singular prefixes. In (36a) we have a 3SG S prefix *d-*, giving a meaning equivalent either to 'he' or 'she', and the same prefix occurs as the 3SG O prefix in (37c), giving rise to the 'him/her' meaning; thus, SO group together. Note that *d-* is only an SO form, and is of course not used to mark a third person singular A, since we're dealing with a grouping of SO vs. A here. Instead, the 3SG A prefixes in (37a) and (37b) occur in an entirely different form, and moreover they're differentiated according to gender (*l-* for the 3SG feminine, and *y-* for the 3SG masculine), which the SO form isn't.

Next, let's look at first person plural prefixes. In (36b) we have a 1PL S prefix *h-*, giving the 'we' meaning, and the same prefix occurs as the 1PL O prefix in (37a) and (37b), giving the 'us' meaning. The data contain an additional complication which you may have noticed: *h-* also means first person plural ('we') in (37c), where it's an A, rather than S or O. How then do native speakers of Abaza know what's going on? The answer is that because the order of prefixes is fixed, the data indicate clearly to an Abaza speaker that *h-* really is the A argument in (37c), the subject of the transitive verb, since it follows the O prefix. It is rather common for languages to 're-use' pieces of morphology in this way: as long as they are clear in their context, duplications of this kind don't appear to cause confusion.

6.4.4 Split systems II

A language with an ergative agreement system may have ergative case marking too (for instance, the North-East Caucasian language Avar) but it is also possible to have ergative agreement on the verb but no case marking on independent NPs – in fact, Abaza (illustrated above) falls into this category. There are also languages which have ergative case marking on NPs, but a nominative/accusative system of cross-referencing on the verb. The Australian language Warlpiri illustrates this system. Full noun phrases and independent (i.e. freestanding) pronouns are all marked with

ergative/absolutive case. The agreement markers (in bold) are affixed to the auxiliary, the element in second position in (38) and (39):

- (38) Ngaju **ka-rna** wangka-mi. (Warlpiri)
 I.ABS AUX.PRES-1SG.SU speak-NONPAST
 ‘I(S) am speaking.’
- (39) Ngajulu-rlu **ka-rna-ngku** nyuntu nya-nyi.
 I-ERG AUX.PRES-1SG.SU-2.OBJ you.ABS see-NONPAST
 ‘I(A) see you(O).’

Look first at the independent pronouns in (38) and (39), which are marked according to the ergative/absolutive system. The S argument *ngaju* (‘I’) in (38) is absolutive, as is the O argument *nyuntu* (‘you’) in (39). Conversely, the A argument *ngajulu-rlu* (‘I’) in (39) is ergative. SO thus group together in opposition to A, as we anticipate in this system. We clearly see that the pronoun for ‘I’ has a different case according to whether it’s the subject of an intransitive verb (S), as in (38), or a transitive verb (A), as in (39).

Now look in contrast at the verb agreement. This marks *both* instances of first person singular in the same way, with the suffix *-rna* designating any first person singular subject. So the affixes reflect a grouping of all subjects (AS) as opposed to all objects (O), namely a nominative/accusative system. The Warlpiri system is not at all unusual, whereas there are no known languages with accusative case systems but ergative agreement systems. This is, then, another way in which the accusative system predominates cross-linguistically.

6.5 GRAMMATICAL RELATIONS

6.5.1 Investigating core grammatical relations

In this section we examine the cross-linguistic properties of two major core grammatical relations, SUBJECT and OBJECT. To show that these concepts are valid, we need to demonstrate that certain linguistic phenomena are best described in terms of ‘subject’ or ‘object’. For languages in the nominative/accusative class, it’s clear that ‘subject’ and ‘object’ are valid categories: in the last few sections we’ve seen a number of illustrations of both case and verb agreement operating in terms of subject vs. object alignment. The examples seen so far show that certain languages are *morphologically* nominative/accusative. This means that the characteristic AS/O alignment is indicated by changes in the morphology (form) of the NPs, via case marking, or in the morphology of the head verb, via agreement, or indeed by marking on both NPs and verbs. However, the AS/O pattern is also pervasive in syntax itself. This means that many languages – including those with no case marking or even with ergative case marking – are *syntactically* nominative/accusative. In such languages there are a number of syntactic processes which

revolve around the subject and object relations – in fact, particularly the subject, since this grammatical relation is by far the most important. We'll examine some of these processes in this section, and return to this topic in Chapter 7.

The subject relation is crucial cross-linguistically: subjects tend to control the syntax in a number of ways, as we'll see. However, it's hard to give a satisfactory definition of 'subject', because no single property is shared by all subjects in all languages. Instead, there's a set of properties typical of subjects, and each language is likely to exhibit a subset of these properties. We begin by looking at some of the main cross-linguistic properties of subjects (Section 6.5.2), and then turn to the question of subjecthood in specific languages (Section 6.5.3).

6.5.2 Subjects: Typical cross-linguistic properties

- i. Subjects are normally used to express the AGENT of the action, if there is an agent.
- ii. Subjects tend to appear first in the clause in unmarked (basic) constituent order. Recall that maybe 80 per cent of languages are either SOV or SVO, therefore subject-initial. But since that leaves many languages that are not subject-initial, we can't use this as a defining property.
- iii. Subjects are understood as the missing argument in IMPERATIVE constructions. An imperative is a command such as *Sit!* or *Eat up your greens!*. Both intransitive and transitive verbs have an understood (or in some languages, overt) second person subject pronoun ('you') in the imperative.
- iv. Subjects control REFLEXIVE NPs, that is, '-self' forms such as the English *herself*, *themselves*, and also RECIPROCAL NPs such as *each other*. So we get *My sister really admires herself*, where the NP *herself* (feminine singular) refers back to the feminine singular subject, *my sister*, but we don't get **Herself really admires my sister*. Note that we can't simply say that the reflexive must refer to a preceding NP. We see this in the Malagasy language Malagasy, which has VOS order, so the subject does not precede the '-self' form. Nonetheless, the subject determines the reference of the '-self' NP; that is, the subject determines which NP the '-self' form refers to.

- (40) a. Manaja tena Rabe. (Malagasy)
 respect self Rabe
 'Rabe respects himself.'
- b. *Manaja an-dRabe tena.
 respect ACC-Rabe self
 '*Himself respects Rabe.'

In (40b), *Rabe* is the object, as we can tell from its accusative case marker; only when *Rabe* is the subject is the reflexive sentence grammatical, as in (40a).

- v. Subjects often control the referential properties of an NP in another clause. For instance, when two clauses are conjoined, as in (41), the subject of the second clause can be omitted because it is CO-REFERENTIAL with the subject of the first clause, *Chris*: I show the omitted NP with \emptyset . But it's only the subject that can be omitted, (41a), not the object, (41b). Moreover, the NP that's omitted has to refer back to the subject of the first clause, *Chris*, and not the object, *Lee*. The subscripts $_i$ and $_j$ here have no meaning of their own, but are simply labels to show which NPs co-refer (= designate the same entity).

- (41) a. [Chris_i phoned Lee_j] and [\emptyset _i met him_j later].
 b. *[Chris_i phoned Lee_j] and [he_i met \emptyset _j later].

Second, in many languages verbs like 'begin' and 'want' take an infinitival complement clause, as in *Kim began [to grate the carrots]*. The 'understood' subject of the 'grate' clause is co-referential with the main clause subject: we understand that it's Kim who is grating the carrots. But only the subject in the infinitival clause – and not the object – can be the 'understood' NP:

- (42) a. Chris_i wants [\emptyset _i to meet this famous film star].
 b. *Chris_i wants [this famous film star to meet \emptyset _i].

- vi. Subjects are the most usual target for promotion from other positions. For instance the PASSIVE construction promotes an NP from direct object position to subject position (see Chapter 7), turning *The students applauded her* into *She was applauded (by the students)*: the pronoun has the form *her* as an object, but *she* as a subject. Although not all languages have promotion processes, if a language has any promotion processes, then it will have ones that move some constituent into subject position.

6.5.3 An examination of subjects in specific languages

We turn now to an examination of subjects in particular languages. We look first at Icelandic (Section 6.5.3.1), which has nominative/accusative morphology and syntax, and so has a clear SUBJECT relation. Section 6.5.3.2 then turns to a morphologically ergative language, Lezgian, for which the notion of subject is more controversial. Section 6.5.3.3 examines Tagalog, which represents a language type different to both accusative and ergative. Section 6.5.3.4 asks whether there are universal grammatical relations.

6.5.3.1 Icelandic

Icelandic is a standard accusative language – subjects are usually in the NOMINATIVE case and objects in the ACCUSATIVE case:

- (43) Ég sá stúlkuna.
 I.NOM saw.1SG the.girl.ACC
 'I saw the girl.'

Verbs in Icelandic agree in person and number with the nominative subject:

- (44) a. Við dönsuðum.
 we.NOM danced.1PL
 ‘We danced.’
 b. Þeir dóu.
 they.NOM died.3PL
 ‘They died.’

However, some verbs are exceptional: their subjects take a case other than nominative. In (45), we have a dative subject, and in (46), an accusative subject, *hana* ‘her’ (the object is also accusative in (46)):

- (45) **Henni** leiddist.
 her.DATIVE bored
 ‘She was bored.’
 (46) **Hana** vantar peninga.
 her.ACC lacks money.ACC
 ‘She lacks money.’

These subjects with ‘quirky’ case don’t trigger subject/verb agreement. In (47), the subject is a *plural* pronoun *þá*, ‘them’ (accusative), but we find the same form of the verb *vantar*, ‘lacks’, as in (46) where the subject is *singular*. Compare (44b), where the nominative *þeir* ‘they’ triggers agreement, giving a plural form of the verb:

- (47) **Þá** vantar peninga.
 them.ACC lacks money.ACC
 ‘They lack money.’

So if these ‘quirky’ subjects don’t trigger verb agreement, on what grounds can we say they’re subjects? There are, in fact, a number of diagnostics for subjects in Icelandic, and the NPs with quirky case pass all of these tests. First, subjects can undergo subject/verb inversion (see Section 3.2.4 on inversion in English). Example (48) shows that an ordinary nominative subject inverts with the finite verb to form a yes/no question, and in (49), we see that a dative subject also inverts. The subjects are in bold type:

- (48) Hafði **Sigga** aldrei hjálpað Haraldi?
 had Sigga.NOM never helped Harold.DATIVE
 ‘Had Sigga never helped Harold?’
 (49) Hefur **henni** alltaf þótt Ólafur leiðinlegur?
 has her.DATIVE always thought Olaf.NOM boring
 ‘Has she always thought Olaf boring?’

Even though there's also a nominative NP *Ólafur* in (49), this couldn't be inverted with the verb *hefur* 'has'.

Second, when two clauses are conjoined, the subject of the second clause can be omitted when it's co-referential with the subject of the first clause, just as in English: see (v) in Section 6.5.2 above. Example (50) illustrates with ordinary nominative subjects:

- (50) Þeir fluttu líkið og (þeir) grófu það.
 they.NOM moved the.corpse and they.NOM buried it
 'They moved the corpse and (they) buried it.'

Turning next to a quirky subject, we see in (51) that the verb meaning 'like' takes a dative subject:

- (51) Mér líkar vel við hana.
 me.DATIVE likes well with her
 'I like her.'

And it turns out that this dative subject can undergo this SUBJECT ELLIPSIS (= omission) too: the dative subject pronoun *mér* can be omitted in the second clause in (52):

- (52) Ég sá stúlkuna og (mér) líkaði vel við hana.
 I.NOM saw the.girl.ACC and me.DATIVE liked well with her
 'I saw the girl and (I) liked her.'

In fact not only can a quirky subject *undergo* ellipsis, as in (52), it can also be the NP which *permits* ellipsis of another subject. This, then, is the third test for subjecthood. Example (53) has a dative subject in the first clause, and the nominative subject *þeir* can undergo ellipsis in the second clause:

- (53) Þeim líkar maturinn og (þeir) borða mikið.
 them.DATIVE likes the.food.NOM and they.NOM eats much
 'They like the food and (they) eat a lot.'

Note that even though there is a nominative NP in the first clause, this is not the subject, and a missing subject can't refer back to it: the sentence couldn't mean, even jokingly, that the food eats a lot.

In sum, then, these (and other) tests for subjecthood in Icelandic show that subject NPs with quirky case really *are* subjects, despite the fact that they fail to trigger subject/verb agreement.

6.5.3.2 Lezgian

We saw earlier that morphologically ERGATIVE languages (i.e. those with ergative case and/or agreement) may be syntactically ACCUSATIVE. This means that syntactic

constructions such as subject ellipsis utilize a grouping of the S and A arguments, as opposed to the O argument ($S = A \neq O$). In fact, it is quite usual for languages which have morphologically ergative alignment to be accusative in terms of their *syntax*, and much rarer for them to have ergative syntax. We will see more on this in Chapter 7.

As we saw in Section 6.3.3, Lezgian is morphologically ergative: the case marking on NPs contrasts absolutive (on S and O noun phrases) with ergative (on A noun phrases, the subjects of transitive verbs): $S = O \neq A$. It will help to review the discussion of (18) through (20) before reading further.

Evidence of syntactic accusativity in Lezgian comes from the fact that it has a SUBJECT grammatical relation (Haspelmath 1993). Let's look first at some basic data. The 'subject' consists of three NP types. The first two types are the A and S arguments: these are the two NPs that would constitute the 'subject' relation in an accusative language. To illustrate these two, we have the ergative-marked argument (A) of a transitive verb, as in (54), and the absolutive-marked argument (S) of an intransitive verb, as in (55). The NPs in bold type in (54) to (56) are the putative subjects.

- (54) **Ruš-a** gadadi-z cük ga-na.
 girl-ERG boy-DATIVE flower.ABS give-PAST
 'The girl gave a flower to the boy.'

- (55) **Ruš** elq̄wena q'uluq^hdi kilig-na.
 girl.ABS turn backward look-PAST
 'The girl turned around and looked back.'

The third potential 'subject' is the experiencer argument of verbs with meanings such as 'want', 'see' and 'be afraid', which in Lezgian take the DATIVE case, as in (56); cross-linguistically, this use of dative case for the semantic role of experiencer is quite common.

- (56) **Ruša-z** ada-q^haj kič'e xa-na-č.
 girl-DATIVE he-of afraid be-PAST-NEG
 'The girl wasn't afraid of him.'

Note that all three putative subject NPs have different cases, so we certainly can't identify 'subjects' by their morphological case in Lezgian. Furthermore, although the NP in bold in (55) is absolutive, not all absolutive NPs are subjects, of course: the O noun phrase *cük* 'flower' in (54) isn't. Similarly, not all dative NPs are subjects: *gadadiz* 'boy' in (54) isn't the subject – dative case here identifies the indirect object (see Section 6.5.4). If there is a 'subject' grammatical relation in Lezgian, then, it cuts across the morphological case marking.

So why would anyone think that Lezgian has a 'subject' relation? Constituent order provides some indication that all three NP types in bold in (54) through (56) pattern together: all have the same clause-initial position, which, as we know from Section 6.2, is the most common position for subjects cross-linguistically. But

position alone won't uniquely identify subjects in Lezgian, because the constituent order is actually very free, so other NP types can be initial in the clause.

However, we can test for subjects using a construction parallel to that in (42) in Section 6.5.2 – please look back to check on this – in which an embedded infinitival clause has an understood subject that refers back to the main clause subject. Look first at the English translations in (57) through (59) to get the idea of the construction, which is very similar in the two languages: the infinitival clause is the complement of a finite verb 'wants' in the matrix clause. The main difference is that in Lezgian, the infinitival clause (shown in square brackets) precedes the finite verb *k'anzawa* 'wants', whilst in English the embedded clause follows *wants*. Crucially, the *understood* subject in Lezgian (marked with \emptyset) can only be one of the three NP types tentatively identified above as forming a 'subject' category: either an ergative subject (an A), an absolutive subject (an S) or a dative subject.

- (57) Nabisata- z_i [[NP_{Erg} \emptyset_i]] ktab k'el-iz] k'an-zawa.
 Nabisat-DATIVE (Subject) book read-INFIN want-IMPF
 'Nabisat wants to read a book.'
- (58) Nabisata- z_i [[NP_{Abs} \emptyset_i]] q'hüre-z] k'an-zawa.
 Nabisat-DATIVE (Subject) laugh-INFIN want-IMPF
 'Nabisat wants to laugh.'
- (59) Nabisata- z_i [[NP_{Dative} \emptyset_i]] xwa akwa-z] k'an-zawa.
 Nabisat-DATIVE (Subject) son see-INFIN want-IMPF
 'Nabisat wants to see her son.'

We know what case the understood subject would have in each example by looking at what happens in ordinary finite clauses with overt (= pronounced) subjects: the verb for 'read' takes an ergative subject, the verb for 'laugh' an absolutive subject, and the verb for 'see' a dative subject. Compare (60): here, the understood NP is again absolutive, but (60) is ungrammatical because this absolutive NP is an O, an absolutive *object* (Musa is the one being sent) rather than an S, an absolutive *subject* as in (58).

- (60) *Musa- z_i [didedi [NP_{Abs} \emptyset_i]] šherdi-z raqur-iz] k'an-zawa.
 Musa-DATIVE mother.ERG (Abs.Object) town-DATIVE send-INFIN want-IMPF
 ('Musa wants to be sent to town by his mother.')

In sum, the Lezgian data show that even a morphologically ergative language may display syntactic accusativity, and indeed there does seem to be evidence for a subject relation in Lezgian.

6.5.3.3 Tagalog

In this section we will examine a language which resists clear classification into either the accusative or the ergative type, and seems in fact to have a totally different marking system for NPs. In Tagalog and other languages of the Philippines, NPs are not case-marked, but they are each preceded by a marker (which we can consider a preposition) that indicates their SEMANTIC ROLE (see Section 2.3.1). The preposition *ng* marks both agent and theme; *sa* (or *mula sa*) marks locative – i.e. indicating location, and glossed as ‘from’ in (61); and *para sa* marks beneficiary, glossed as ‘for’ in (61). However, in every sentence *one* of the NP participants must be chosen to be the TOPIC of the clause, and it is marked as such by a special preposition, *ang*, which replaces the marker the NP would have otherwise. The topic is shown in bold in each example. (Note that the topic is always understood to be definite, whilst the other NPs can be understood as definite or indefinite.)

Also, the verb itself has an affix that marks the semantic role of the NP chosen as topic: I’ve indicated this role beneath the gloss for the verb in each example. This marking is clearly a kind of verb agreement, but it is different from either the accusative system or the ergative system in that it does not operate in terms of the grammatical function of the NP arguments. In examples like (61), any one of the NP participants can be marked as the topic – and whichever semantic role the topic has will be indicated on the verb, resulting in a verb marked to agree with one of the properties ‘agent’, ‘theme’, ‘locative’ or ‘beneficiary’; this is shown in the different morphology that the verb has in each example in (61). To see this, you’ll need to study the Tagalog data itself, as the gloss doesn’t reflect the morphological distinctions.

- (61) a. Kukuha **ang babae** ng bigas sa sako para sa bata.
 FUT.take.out TOPIC woman THEME rice from sack for child
 AGENT.TOPIC
 ‘The woman will take some rice out of a sack for a/the child.’
- b. Kukunin ng babae **ang bigas** sa sako para sa bata.
 FUT.take.out AGENT woman TOPIC rice from sack for child
 THEME.TOPIC
 ‘A/the woman will take the rice out of a sack for a/the child.’
- c. **Ang sako** ay kukunan ng bigas ng babae para sa bata.
 TOPIC sack be FUT.take.out THEME rice AGENT woman for child
 LOCATIVE.TOPIC
 ‘The sack will have rice taken out of it by the woman for the child.’
- d. **Ang bata** ay ikukuha ng bigas ng babae mula sa sako.
 TOPIC child be FUT.take.out THEME rice AGENT woman from sack
 BENEFICIARY.TOPIC
 ‘The child will have rice taken out of the sack for him/her by the woman.’

It's clear, then, that Tagalog isn't morphologically marked in accordance with either the accusative system or the ergative system, either by case marking or by verbal agreement.

However, as we have already noted, a language may nonetheless be *syntactically* accusative despite not being morphologically accusative. Does Tagalog fit this pattern? Looking at the syntactic behaviour of NPs, it turns out that some processes operate in terms of topics, irrespective of their semantic and syntactic role. But there are also other processes that operate in terms of a grouping of the A and S noun phrases, whether or not they are topics: this is a syntactically accusative pattern, and suggests that there may after all be a 'subject' in Tagalog.

Let's look first at a process that targets topics: the 'all' construction. In (62) we see that *lahat*, 'all', is understood as modifying whichever NP is the topic. In (62a) the topic is the agent, so *lahat* must modify the A noun phrase, meaning 'the children'. But in (62b) we have a theme topic, referring to the 'thing written', so *lahat* must modify the O noun phrase, meaning 'the letters'. Note that *lahat* is not even adjacent to this latter phrase, *ang mga liham*, in (62b):

- (62) a. Susulat lahat **ang mga bata** ng mga liham.
 FUT.write all TOPIC PL child THEME PL letter
AGENT.TOPIC
 'All the children will write letters.'
- b. Susulatin lahat ng mga bata **ang mga liham**.
 FUT.write all AGENT PL child TOPIC PL letter
THEME.TOPIC
 'The/some children will write all the letters.'
not 'All the children will write letters.'

Since linguists don't normally consider a grouping of A and O to form any grammatical relation, the 'all' construction favours a view of Tagalog as not having subjects.

Now consider a syntactic process in Tagalog that targets S and A noun phrases – the classic 'subject' pairing – and not topics. Our examples are complement clauses with an understood subject. You will find it helpful to look again at the discussion of similar examples given earlier on before reading further: see (42) from English and (57) through (59) from Lezgian. In the Tagalog construction, the 'missing' subject is always an S or an A, whether or not it's a topic. In the examples in (63), the topic of each matrix clause (the 'hesitate' clause) is the agent, *siya*, meaning 'he' (the suffix *-ng*, which I have left unglossed, indicates that an embedded clause follows). In the embedded clause, however, the topic is different in each example:

- (63) a. Nagatubili siya-ng [humiram ng pera sa banko].
 hesitate he.TOPIC-NG borrow THEME money from bank
AGENT.TOPIC *AGENT.TOPIC*
 'He hesitated to borrow money from a/the bank.'

- b. Nagatubili siya-ng [hiramin **ang pera** sa banko].
 hesitate he.TOPIC-NG borrow TOPIC money from bank
AGENT.TOPIC *THEME.TOPIC*
 ‘He hesitated to borrow **the money** from the bank.’

Remember that these ‘borrow’ clauses are embedded clauses with an understood subject. In both embedded clauses in (63a) and (63b), this understood subject is the agent (referring back to ‘he’ in the matrix clause), an A noun phrase. In (63a), the missing agent NP in the ‘borrow’ clause happens also to be the NP chosen as the topic, as we can tell from the form of the verb, which, as you’ll recall, is marked for the semantic role of the topic. In (63a), then, there isn’t an overt *ang*-NP – an overt topic – because this topic is the ‘understood’ NP. But in (63b), the topic of the embedded clause is the theme (thing borrowed), namely *ang pera* ‘the money’, yet the understood subject of that clause is still the A noun phrase (referring back to ‘he’). So topics are clearly *not* the targets for the ellipsis (the omitted part) in this construction. In fact, this process of NP ellipsis suggests that Tagalog is syntactically accusative, at least in this one construction. In other words, the understood subject can be either an S or an A argument (i.e. any type of ‘subject’), but not an O argument. Any process that treats S and A noun phrases together – and O arguments differently – suggests that the language operates at least part of its syntax in terms of a nominative/accusative alignment. Tagalog may indeed, then, have a SUBJECT grammatical relation consisting of S and A.

6.5.3.4 *Language universals?*

We are left with the indication that the ‘subject’ relation is very important cross-linguistically, even occurring in some languages which are not otherwise nominative/accusative in their morphology and elsewhere in their syntax. Should it be considered a language universal? Some linguists argue that it should not. For instance, Dryer (1997) argues that there are no universal grammatical relations, discussing data from Dyirbal, Cree, Cebuano and Acehnese which are particularly problematic for a view that ‘subject’ is a universal category. This view is supported by Croft (2001); see also Haspelmath (2007). We will leave this an open question. However, it is clear that recurring properties in grammatical relations *are* found cross-linguistically, among languages of very different syntactic types and from totally unrelated language stocks. It seems, then, that these properties may reflect some universal features, even if these are not fully understood at present.

Finally, we saw in Section 6.5 that languages which have ergative (or no) case marking may nonetheless exhibit nominative/accusative syntax. Is the opposite situation ever seen – in other words, can a language be syntactically ergative even if it doesn’t have ergative/absolutive case morphology? Linguists used to think not, but more recently it has been shown that this situation does exist (Donohue and Brown 1999). Certainly, though, it is very rare.

6.5.4 Objects

The other major grammatical relation is that of **OBJECT**, in accusative systems the complement of a two-argument verb. There is plenty of morphological evidence for the existence of an object relation in languages with nominative/accusative morphology, since the O argument is designated by a special case (accusative) and/or verb agreement. This chapter contains examples of a case-marked O from languages as genetically diverse as Latin (1), Turkish (4) and Dyirbal (25). (Recall from Section 6.3.4 that Dyirbal is largely ergative, but its first/second person pronouns have an accusative case system.) Verb agreement with the O argument is shown in several examples: see (30) from Kambera and (39) from Warlpiri.

Syntactic evidence for the O relation is more limited than for subjects, but in many languages only an O can be passivized (see 1.1.1, 1.3.2 and also Chapter 7 for a demonstration of this). Recall that in Icelandic, we find certain constructions in which a noun phrase doesn't receive the expected case marking, but instead gets a 'quirky' case: examples of *subject* NPs with quirky case were given in Section 6.5.3.1. Icelandic also has certain *object* NPs with quirky case, so we can see how these act in terms of typical object behaviour. It turns out that not only do ordinary accusative O arguments undergo passivization, so too do O arguments with quirky case. An example of a quirky object is the NP *mér* in (64): this is not accusative, as objects typically are in Icelandic, but rather it is dative:

- (64) Þeir hjálpuðu mér. (Icelandic)
 they.NOM helped me.DATIVE
 'They helped me.'

Like other objects, however, this O can be promoted to subject position, giving (65). Note, though, that the dative case remains on this NP – it doesn't become nominative – though its position is the standard clause-initial position of the subject in Icelandic:

- (65) Mér var hjálpað. (Icelandic)
 me.DATIVE was helped
 'I was helped.'

Verbs such as 'give', 'send' and 'show', which take three arguments (*X gave Y to Z*), can in some languages be said to distinguish a **DIRECT OBJECT** from an **INDIRECT OBJECT**. In accusative languages with extensive case systems, the direct object bears accusative case, whilst what is traditionally termed the indirect object bears **DATIVE** case, as in Turkish, German, Greek and Latin. The indirect object is typically the 'recipient' or 'goal' NP, such as *mir* in (66):

- (66) Mein Freund gab mir sein Fahrrad. (German)
 my.NOM friend gave me.DATIVE his.ACC bicycle
 'My friend gave me his bicycle.'

The dative is also used for this same purpose in many ergative languages: see (54) from Lezgian. Cross-linguistically, then, the central use of the dative case is to designate the NP that's the recipient or the beneficiary or the goal of a three-argument verb.

But this type of NP does not always get a special case. For instance, although in Ancient Greek most three-argument verbs have an accusative direct object and a dative indirect object, the verb for 'teach' is exceptional in that both of its complements (the NPs meaning 'the boy' and 'the music') have accusative case:

- (67) Edidaxan [ton paida] [tēn mousikēn]. (Ancient Greek)
 taught.3PL the.ACC boy.ACC the.ACC music.ACC
 'They taught the boy music.'

In fact, in English and many other languages there is little justification for distinguishing an 'indirect object' from any other object. Very often, the recipient NP looks just like a direct object – in what is known as the DOUBLE OBJECT construction, the recipient immediately follows the verb and has the same case marking as any object, as in *Kim lent **me** the book*. Alternatively, the recipient appears in an ordinary PP headed by 'to' or 'for', as in *Kim made a cake **for me***. In other words, there's neither a special case nor any special syntactic behaviour associated with the NP that traditional grammar calls the indirect object.

In Section 6.4 we saw that a ditransitive verb such as 'give' may agree with all three of its argument NPs – see (35) from Biri. However, a more common situation is that only two arguments of a three-argument verb are actually marked on the verb. One is always the subject, but languages differ in terms of which other NP the verb agrees with: it can be either the NP with the semantic role of theme (such as 'thing given'), or else the recipient. Commonly, the verb agrees with the *recipient* NP, rather than the 'thing given'. Example (68) illustrates from Warlpiri, which, as we saw in (38) and (39) above, has ergative case marking but accusative verb agreement:

- (68) Ngaju-ku ka-**npa**-ju karli yi-nyi nyuntulu-rlu.
 me-DATIVE PRES-2SG.SU-1SG.OBJ boomerang give-NONPAST you-ERG
 'You are giving me a boomerang.'

There are two agreement suffixes in (68). The first, *-npa*, marks the subject ('you'). There is no agreement marker for *karli*, 'boomerang', on the verb at all, but the second suffix is a 1SG marker which cross-references the recipient, the dative NP *ngajuku*, meaning 'me'. What's more, the 1SG agreement suffix *-ju* is the same suffix that is used to mark the first person singular O argument of an ordinary transitive verb (as in 'You saw me'). It appears, then, that in some languages the recipient functions as a kind of object. Perhaps it's not too surprising that the verb marks the recipient here: recipients are typically human, or at least animate, and thus arguably more important than an inanimate theme NP such as *karli*, 'boomerang'.

6.6 FREE WORD ORDER: A CASE STUDY

Having examined case, agreement and grammatical relations, we are now in a position to return to the topic of word order. We have already seen that languages with extensive case marking on noun phrases typically allow much variation in constituent order (see, for example, the German data in Section 4.3.3.1, and also the discussion of Japanese in Section 8.3). The same is true of languages with extensive head-marking on the verb, such as Kambera. The current section shows that some languages also allow extremely free *word* order, in the most literal sense. One such language is Latin; another is Navajo. Our illustrations, though, are from Australian languages, and in particular, Warlpiri. To remind you, Warlpiri exhibits a split ergative system: it has ergative/absolutive case for independent noun phrases and pronouns, but an accusative system for the pronominal affixes marked on the auxiliary (Section 6.4.4 above).

First, we illustrate the fact that Warlpiri has free *constituent* order: the only restriction is that the auxiliary, expressing tense and person/number marking, must be in second position in the clause.

- (69) a. Ngarrka-ngku ka wawirri panti-rni. (Warlpiri)
 man-ERG AUX.PRES kangaroo.ABS spear-NONPAST
 'The man is spearing the kangaroo.'
- b. Wawirri ka panti-rni ngarrka-ngku.
 kangaroo.ABS AUX.PRES spear-NONPAST man-ERG
 'The man is spearing the kangaroo.'
- c. Panti-rni ka ngarrka-ngku wawirri.
 spear-NONPAST AUX.PRES man-ERG kangaroo.ABS
 'The man is spearing the kangaroo.'

These three as well as the other three orders of S, O and V are all possible, with no single basic order. Two further Warlpiri examples are shown in (70):

- (70) a. Jarntu-jarra-rlu lpa-pala-jana ngaya nya-ngu.
 dog-DUAL-ERG AUX.IMPF-3.DUAL.SU-3PL.OBJ cat.ABS see- PAST
 'The two dogs were looking at the cats.'
- b. Ngaya lpa-pala-jana jarntu-jarra-rlu nya-ngu.
 cat.ABS AUX.IMPF-3.DUAL.SU-3PL.OBJ dog-DUAL-ERG see-PAST
 'The two dogs were looking at the cats.'

These two sentences mean the same thing, and indeed would be considered the same sentence by a Warlpiri speaker (Shopen 2001: 191). Hale (1983: 5) reports that 'different linear arrangements count as repetitions of each other' in Warlpiri. However, Shopen notes that moving an element to clause-initial position in Warlpiri signals its pragmatic importance, making the initial element the focus or topic of the sentence.

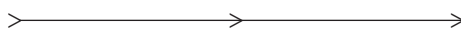
Next we see that constituents can also be split up in Warlpiri, so that the word order is literally free: the auxiliary must still be *either* the second constituent (71a), *or* second word (71b), but this remains the only restriction. (Actually, the auxiliary attaches to the end of the first constituent or first word in the clause, though the notation here doesn't show that.) So a noun phrase such as *wawirri yalumpu*, 'that kangaroo', can appear either as in (71a) or as in (71b):

- (71) a. **Wawirri** **yalumpu** kapi-rna panti-rni. (Warlpiri)
 kangaroo.ABS that.ABS AUX.FUT-1SG.SU spear-NONPAST
 'I will spear that kangaroo.'
- b. **Wawirri** kapi-rna panti-rni **yalumpu**.
 kangaroo.ABS AUX.FUT-1SG.SU spear-NONPAST that.ABS
 'I will spear that kangaroo.'

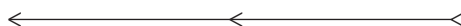
In (71b) we have a DISCONTINUOUS CONSTITUENT; the elements of the absolutive O noun phrase in bold type are not contiguous. It's possible for the O argument to be freely split up in this way because the case marking identifies its components as belonging to the same, absolutive NP (though, quite typically for absolutive case, there is no overt case suffix here).

Let's now turn to some more complex examples. In (72a), there's a continuous A constituent, *maliki wiringki* 'big dog'. But in (72b) and (c), the individual elements of this constituent are not contiguous, but instead are split up in two different ways:

- (72) a. **Maliki** **wiri-ngki** Ø-ji yarlku-rnu. (Warlpiri)
 dog big-ERG AUX.PAST-1SG.OBJ bite-PAST
 'The/a big dog bit me.'
- b. **Maliki-rli** Ø-ji yarlku-rnu **wiri-ngki**.
 dog-ERG AUX.PAST-1SG.OBJ bite-PAST big-ERG
 'The/a big dog bit me.'
- c. **Wiri-ngki** Ø-ji yarlku-rnu **maliki-rli**.
 big-ERG AUX.PAST-1SG.OBJ bite-PAST dog-ERG
 'The/a big dog bit me.'



Before reading further, please examine the examples in (72) and indicate what the difference is between the (a) sentence on the one hand and the (b) and (c) sentences on the other. Use the correct terminology to describe this. Why might this difference occur, do you think?



In (72a), there is only one ergative case marker, the suffix *-ngki*, and it's attached to the end of the whole the A argument, which here is a continuous NP. But in both (72b) and (72c), each element of the discontinuous NP has an ergative suffix. It

wouldn't be ungrammatical to use an ergative suffix on *maliki* in (72a) too; but crucially, that suffix can't be omitted in (72b) and (72c), where the NP is discontinuous. Again, the case marking identifies each subpart of the discontinuous A argument.

In (71) and (72), only one NP is discontinuous. However, in free word order languages it's also perfectly possible to have, say, both of the arguments of a transitive verb as split NPs. This example is from another Australian language, Kalkatungu, which 'exhibits a marked tendency to represent noun phrases discontinuously' (Blake 2001b: 419). Here, the two discontinuous NPs are interleaved:

- (73) **Tjipa-yi** tjaa **kunka-ngku** pukutjurrka lhayi **nguyi-nyin-tu**. (Kalkatungu)
 this-ERG this branch-ERG mouse kill fall-PARTICIPLE-ERG
 'The falling branch killed the mouse.'

The elements in bold type are the subparts of the ergative A argument which means 'the falling branch', and each has an overt case suffix which identifies them as ergative. The underlined elements *tjaa pukutjurrka* form the O argument, and these receive no overt case marking; this is, of course, superfluous, since the A argument is already marked. Thus, having affixes just on the subparts of one NP is enough to ensure that there's no ambiguity. Once again, I stress that discontinuous phrases in these languages are by no means exceptional – quite the opposite, in fact.

We might wonder whether, in free word order languages, we should use the term 'constituent'; after all, if noun phrases can be split up so readily, is it appropriate to describe the syntax of these languages in terms of 'constituency' at all? It has sometimes been claimed not (e.g. Evans and Levinson 2009). Instead, it may be more appropriate to describe the syntax solely in terms of DEPENDENCIES, so that in examples like (71) to (73), what really counts is the word-to-word relationships, as indicated, for instance, by the case markings on each related element.

Nonetheless, it seems that constituent structure does play a role in free word order languages. In Warlpiri examples like (72), sentences with continuous NP arguments don't have just the same range of meanings as those with discontinuous constituents (Hale 1983; Austin and Bresnan 1996). The discontinuous constituents in (72b) and (72c) give rise to an additional meaning, which is 'The/a dog bit me and it was big.' But the continuous NP in (72a) has the 'merged' meaning: it can *only* mean 'The/a big dog bit me'. This distinction clearly suggests that NP constituents really do exist in the language. In addition, we saw that Warlpiri requires the auxiliary to be in second position in the clause: what precedes it can be an NP (71a), or a single word of some kind, including a noun (71b), a verb (69c), and a particle. Crucially, a random sequence of words which don't form a constituent *cannot* precede the auxiliary; this, then, is one test for constituent structure in Warlpiri. If constituents have a different syntactic status to random strings of words, this again suggests that constituents are real in such languages. Finally, examples like (72) showed that in Warlpiri, only the final element in a continuous NP constituent needs to be case marked – (72a) vs. (72b/c). This indicates that there truly is an NP

in (72a): it acts as a unit, so each of its subparts doesn't need a case affix. It seems clear, then, that constituent structure does play a vital role even in free word order languages.

Interestingly, there is apparently no correlation amongst Australian languages between the existence of discontinuous NPs and the appearance of free constituent order; for instance, Austin and Bresnan (1996) report that Diyari has discontinuous NPs but prefers a fixed SOV constituent order. Moreover, though Warlpiri has an extensive system of bound pronominal marking in the auxiliary (Section 6.4.4), it also appears that, cross-linguistically, this is not a necessary condition for the appearance of split NPs: case marking on independent NPs and pronouns is enough to allow for discontinuous constituents. So for instance, the Australian languages Jiwari, Dyirbal and Yidiny all have discontinuous NPs but lack the pronominal cross-referencing affixes that characterize the Warlpiri auxiliary.

I hope to have shown definitively in this final section that if we only looked at English and its close relatives, we'd be missing out on a great deal of knowledge about the potential of the human language faculty. I also hope that by this point, you're feeling more comfortable about analysing examples from 'exotic' languages. More are to come in Chapter 7!

6.7 SUMMARY

This chapter has examined three different ways in which languages represent the relationships between core NPs and the verbal predicate on which they are dependent: constituent order, case marking and verb agreement. All languages use at least one of these methods, and often more than one. Constituent order may be very free or very fixed. In languages with free constituent order (or word order), it is more likely that there will be some system of either dependent-marking (case) or head-marking (agreement) in order to identify the grammatical relation of each core NP participant. The two main case systems are the accusative and the ergative systems. Some languages, such as Chinese, have neither case nor agreement. But even in languages without morphological case, the need to recognize grammatical relations is evident in the syntax. Syntactic constructions generally follow either an ergative or an accusative pattern, the accusative being by far the most common.

Checklist for Chapter 6

If you are uncertain about any of these points, I recommend revising before moving on.

- What are the three major ways in which languages can indicate grammatical relations within the clause?
- What single-letter terms are used to indicate the three main NP functions, and what grammatical relation does each of these represent? Give an example of each.

- What are the two major alignment systems, which group the three main NP functions in two distinct ways?
- What are some of the typical properties of subjects, cross-linguistically?
- What does it mean for a language to have discontinuous constituents?

FURTHER READING

Good places to start on the topics of constituent order, case and agreement would be T. Payne (2006) and Whaley (1997), moving on to Comrie (1989: ch. 4 and ch. 6). The seminal work on constituent order and word order is Greenberg (1966). More recent proposals can be found in Hawkins (1983) and in Tomlin (1986); see also the large body of work by Matthew Dryer, for instance Dryer (1991). On case, see Blake (2001a). On grammatical relations, see Palmer (1994) and the collection of papers in Aikhenvald et al. (2001). On agreement, see Corbett (2006). All of these are textbook treatments and are much recommended. On ergativity, Dixon (1994) is a more advanced read, but absolutely central and very worthwhile. The properties of subjects in Section 6.5.2 are largely taken from Keenan (1976); see also Comrie (1989: ch. 5). The Warlpiri data are largely taken from the work of Ken Hale, a brilliant linguist who undertook extensive fieldwork on endangered languages.

EXERCISES

1. Examine the data in (1) through (3) below (all taken from Stucky 1983). These are simple sentences from the Bantu language Makhuwa, spoken in Tanzania and Mozambique, and they show that the order of phrases is very free in this language. Makhuwa marks both subject and object with cross-referencing agreement prefixes on the verb. The *APPLIC* suffix on the verb is an ‘applicative’ marker; it’s this that gives the sense of preparing porridge for someone, rather than an actual preposition meaning ‘for’, which marks the recipient in the English. This construction is discussed further in Chapter 7.

- (1) Aaarima aho-n-ruw-el-a mwaana isima.
 Aaarima SU-OBJ-prepare-APPLIC-PAST child porridge
 ‘Aaarima prepared porridge for a child.’

- (2) Isima Aaarima aho-n-ruw-el-a mwaana.
 porridge Aaarima SU-OBJ-prepare-APPLIC-PAST child
 ‘Aaarima prepared porridge for a child.’

- (3) Aho-n-ruw-el-a Araarima mwaana isima.
 SU-OBJ-prepare-APPLIC-PAST Araarima child porridge
 'Araarima prepared porridge for a child.'

Each sentence contains four phrases – a subject, a verb, a direct object and an indirect object – but they appear in a different order. In fact, any of the 24 (!) possible orders of the four phrases can be used, given the right context.

Now consider complex sentences: given a subject, a verb and an embedded clause, there are six logically possible orders of these three phrases. However, only three out of the potential six orders are grammatical. The orders actually found in Makhuwa are:

- Subject-verb-embedded clause (4),
- Verb-embedded clause-subject (5),
- and in addition, verb-subject-embedded clause (I haven't illustrated this, but you should be able to reconstruct it).

- (4) Araarima aheew-a [wiira nt'u aho-thek-a iluwani].
 Araarima SU.hear-PAST that someone SU.build-PAST fence
 'Araarima has heard that someone built a fence.'

- (5) Aheew-a [wiira nt'u aho-thek-a iluwani] Araarima.
 SU.hear-PAST that someone SU.build-PAST fence Araarima
 'Araarima has heard that someone built a fence.'

Task: Work out what the three unattested (= non-occurring) phrase orders are and state the generalization about possible phrase orders in Makhuwa. In order to do this, you'll need to look at what the three attested orders and then the three unattested orders have in common. Why might a language have such a restriction, do you think?

2. Examine the data below (slightly adapted from Van Valin 1985) from Lakhota (a native American language, specifically a Siouan language of South Dakota, Montana and Manitoba) and answer questions (i) through (iv).
 - (i) Which argument(s) of the verb, if any, does the verb agree with?
 - (ii) How is agreement (or cross-referencing) indicated in Lakhota? Give the details.
 - (iii) Using the data in (1) through (3) as comparison, try to figure out why (4) and (5) are grammatical, but (6) is ungrammatical. The notation '≠' indicates that the Lakhota form is not a possible way of translating the English sentence given.
 - (iv) In light of your answer to (iii), why do you think (7) is ungrammatical? What generalization can be made about the grammatical vs. ungrammatical examples?

- (1) wičháša ki mathó wə Ø-Ø-kté
man the bear a 3SG.OBJ-3SG.SU-kill
'The man killed a bear.'
- (2) mathó wə wičháša ki Ø-Ø-kté
bear a man the 3SG.OBJ-3SG.SU-kill
'A bear killed the man.'
- (3) wičháša ki mathó óta wičhá-Ø-kté
man the bear many 3PL.OBJ-3SG.SU-kill
'The man killed many bears.'
- (4) wičháša ki ixʔé óta Ø-yəke
man the rock many 3SG.SU-see
'The man saw many rocks.'
- (5) wičháša ki mathó óta wíchá-Ø-yəke
man the bear many 3PL.OBJ-3SG.SU-see
'The man saw many bears.'
- (6) *wičháša ki ixʔé óta wíchá-Ø-yəke
man the rock many 3PL.OBJ-3SG.SU-see
(≠'The man saw many rocks.')
- (7) *ixʔé ki hená hokšíla wə Ø-pi-phá
rock the those boy a 3SG.OBJ-3PL.SU-hit
(≠'Those rocks hit a boy.')

3. In Welsh, the verb agrees with one of its argument NPs, but the conditions on this agreement are somewhat different than in more familiar European languages such as English, French or German. Study the following data, and answer these questions.

- (i) Which NP argument does the verb agree with in Welsh? (Name its GRAMMATICAL RELATION.)
- (ii) What MORPHOSYNTACTIC CATEGORIES of the NP does the verb agree with?
- (iii) What are the restrictions on this agreement?
- (iv) Why are (3), (4) and (8) ungrammatical? Why is the starred alternative in (9) ungrammatical?
- (v) How could you change (4) to make it grammatical, while retaining the meaning?

Hints:

- All data given here are entirely regular, and no data are missing. You have enough information to answer the questions without having to make guesses.

- Welsh has VSO (verb-subject-object) word order, but this is not relevant to your answer.

- (1) Gwelodd y bachgen ddreigiau.
see.PAST.3SG the boy dragons
'The boy saw dragons.'
- (2) Gwelodd y bechgyn ddreigiau.
see.PAST.3SG the boys dragons
'The boys saw dragons.'
- (3) *Gwelson y bechgyn ddreigiau.
see.PAST.3PL the boys dragons
(≠ 'The boys saw dragons.')
- (4) *Gwelson ein ffrindiau ddreigiau.
see.PAST.3PL our friends dragons
(≠ 'Our friends saw dragons.')
- (5) Gwelais i ddreigiau.
see.PAST.1SG I dragons
'I saw dragons.'
- (6) Gwelodd hi / o ddreigiau.
see.PAST.3SG she / he dragons
'She / he saw dragons.'
- (7) Gwelson nhw ddreigiau.
see.PAST.3PL they dragons
'They saw dragons.'
- (8) *Gwelodd nhw ddreigiau.
see.PAST.3SG they dragons
(≠ 'They saw dragons.')
- (9) Aeth / *Aethon y bechgyn allan.
go.PAST.3SG / go.PAST.3PL the boys out
'The boys went out.'
- (10) Aethon nhw allan.
go.PAST.3PL they out
'They went out.'

4. Examine the following data in (1) through (3) (from Blake 1977) and determine what case system is found in the Australian language Yalarnnga: either

nominative/accusative alignment ($S = A \neq O$) or ergative/absolutive alignment ($S = O \neq A$). Make clear what the evidence is for your conclusion. I have indicated the different case markers on the NPs in the gloss by marking one case with **X** and the other with **Y**.

- (1) ngia wakamu
I.X fell
'I fell.'
- (2) kupi-ngku ngia tacamu
fish-Y I.X bit
'A fish bit me.'
- (3) nga-tu kupi-Ø walamu
I-Y fish-X killed
'I killed a fish.'

5. Examine the data from Swahili in (1) through (4) below (from Dixon 1994).

Task: (i) Determine first whether verbal cross-referencing agreement in this language represents a nominative/accusative system or an ergative/absolutive system. Make clear what the evidence is for your conclusion.

Hint:

Remember that the same form may sometimes be used for marking a particular person/number combination in more than one case, as in the Abaza data in (36) and (37) in the text of Chapter 6.

- (1) tu-li-anguka
1PL-PAST-fall
'We fell down.'
- (2) m-li-anguka
2PL-PAST-fall
'You all fell down.'
- (3) m-li-tu-ona
2PL-PAST-1PL-see
'You all saw us.'
- (4) tu-li-wa-ona
1PL-PAST-2PL-see
'We saw you all.'

(ii) Now describe the position of the agreement affixes in Swahili as concisely and accurately as you can, using the correct terminology.

6. Examine the data in (1) through (6) below (from Anderson 1976 and Otsuka 2005) and determine what case system is found in Tongan, either nominative/accusative alignment ($S = A \neq O$) or ergative/absolutive alignment ($S = O \neq A$). Make clear exactly what the evidence is for your conclusion. Finally, describe precisely how case is represented in Tongan. In this exercise, I've simply indicated all the case markers with the *same* gloss, CASE. Of course, they are not all the same, and different markers have different roles in the clause. You will need to work out for yourself which is which, by figuring out the role of each marker! (The character that looks like a quotation mark is a letter of the alphabet in Tongan, and represents a specific consonant, a glottal stop. It has no relevance to the answer here.)

- (1) na'e lea 'a e talavou
PAST speak CASE the young.man
'The young man spoke.'
- (2) na'e ma'u 'e sione 'a e ika
PAST get CASE Sione CASE the fish
'Sione got the fish.'
- (3) na'e alu 'a tevita ki fisi
PAST go CASE David to Fiji
'David went to Fiji.'
- (4) na'e tamate'i 'a kolaiate 'e tevita
PAST kill CASE Goliath CASE David
'David killed Goliath.'
- (5) na'e ma'u 'e siale 'a e me'a'ofa
PAST get CASE Charlie CASE the gift
'Charlie received the gift.'
- (6) na'e kai 'a e ika 'e sione
PAST eat CASE the fish CASE Sione
'Sione ate the fish.'

7. This exercise asks you to investigate patterns of verb agreement in Standard Arabic (Aoun et al. 2010).

Task: Work out the system of subject-verb agreement in the data shown in (1) to (15) below. Account for the grammaticality patterns in all the data given: your answer must cover the ungrammatical as well as the grammatical data. Generalize where possible, so that you produce the most concise account of the agreement facts that is consistent with the data.

Hints:

- Arabic typically allows two distinct constituent orders within the clause: SVO and VSO, though some other word orders occur too.
- Patterns of verbal agreement and pronominals in Arabic reflect distinctions between masculine and feminine gender, but this is not inherently relevant for your account.
- Arabic is a pro-drop or NULL SUBJECT language (Section 1.2), which is why the subject pronouns are shown in parentheses in some of the examples. This means that they are optional in the positions shown.
- The verbs in (1) to (12) are perfective (translated as past tense); the verbs in (13) to (15) are imperfective (translated as present tense). Again, this is not inherently relevant for your account.

(1) ʔakal-at l-muʕallimaat-u
 ate-3F.SG the-teacher.F.PL-NOM
 ‘The (female) teachers ate.’

(2) l-muʕallimaat-u ʔakal-na
 the-teacher.F.PL-NOM ate-3F.PL
 ‘The (female) teachers ate.’

(3) ʔakala l-muʕallim-uun
 ate.3M.SG the-teacher.M.PL-NOM
 ‘The (male) teachers ate.’

(4) *ʔakal-uu l-muʕallim-uun
 ate-3M.PL the-teacher.M.PL-NOM
 (‘The (male) teachers ate.’)

(5) l-muʕallim-uun ʔakal-uu
 the-teacher.M.PL-NOM ate-3M.PL
 ‘The (male) teachers ate.’

(6) *l-muʕallim-uun ʔakala
 the-teacher.M.PL-NOM ate-3M.SG
 (‘The (male) teachers ate.’)

(7) kataba r-risaalat-a l-ʔawlaad-u
 wrote.3M.SG the-letter-ACC the-children-NOM
 ‘The children wrote the letter.’

(8) (hum) katab-uu l-kitaab-a ?ams
 they.M wrote-3M.PL the-book-ACC yesterday
 ‘They (male) wrote the book yesterday.’

(9) (hum) qara?-uu d-dars-a
 they.M read-3M.PL the-lesson-ACC
 ‘They (male) read the lesson.’

(10) qara?-u (humu) d-dars-a
 read-3M.PL they.M the-lesson-ACC
 ‘They (male) read the lesson.’

(11) *qara?a (humu) d-dars-a
 read.3M.SG they.M the-lesson-ACC
 (‘They (male) read the lesson.’)

(12) *hum qara?a d-dars-a
 they.M read.3M.SG the-lesson-ACC
 (‘They (male) read the lesson.’)

(13) ta-drusu T-Taalibaat-u
 3F.SG-study the-students.F.PL-NOM
 ‘The (female) students study.’

(14) T-Taalibaat-u ya-drus-na
 the-students.F.PL-NOM 3-study-F.PL
 ‘The (female) students study.’

(15) ya-drus-na (hun)
 3-study-F.PL they.F
 ‘They (female) study.’

8. For this exercise it will help you to revise Sections 6.5.2 and 6.5.3, on subjects. Recall from this section that in Icelandic, the subjects of some verbs take what is known as ‘quirky’ case. Now examine the Icelandic data in (1) through (3) (the data and arguments on which this exercise is based are from Sigurðsson 1991). You will see that the QUANTIFIER in bold type meaning ‘all’ (which ‘quantifies’ the number of boys) agrees in case with the subject of the clause, as well as in number (plural, here) and gender (masculine, here):

(1) Strákarnir komust **allir** í skóla.
 the.boys.NOM got all.NOM.PL.M to school
 ‘The boys all managed to get to school.’

- (2) Strákana vantaði **alla**.
 the.boys.ACC lacked all.ACC.PL.M
 ‘The boys were all absent.’

- (3) Strákunum leiddist **öllum**.
 the.boys.DATIVE bored all.DATIVE.PL.M
 ‘The boys were all bored.’

Next, examine the data in (4) through (6). These examples are parallel to the construction from Lezgian discussed in Section 6.5.3, and it will help you to revise this particular section.

Task: (i) How can we account for the case marking (as well as the number and gender marking) found on the quantifier meaning ‘all’ in each of the examples in (4) through (6)? (ii) What does the quantifier agree with?

- (4) Strákarnir vonast til [að komast **allir** í skóla.]
 the.boys.NOM hope for to get all.NOM.PL.M to school
 ‘The boys hope to all get to school.’
- (5) Strákarnir vonast til [að vanta ekki **alla** í skólann.]
 the.boys.NOM hope for to lack not all.ACC.PL.M to the.school
 ‘The boys hope to not all be absent from school.’
- (6) Strákarnir vonast til [að leiðast ekki **öllum** í skóla.]
 the.boys.NOM hope for to bore not all.DATIVE.PL.M to school
 ‘The boys hope to not all be bored in school.’