Language Universals and Linguistic Typology

SYNTAX AND MORPHOLOGY

Bernard Comrie

Second Edition

The University of Chicago Press

7

RELATIVE CLAUSES

7.1 SOME TYPOLOGICAL CHARACTERISTICS OF ENGLISH RELATIVE CLAUSES

In this section, the aim is to indicate some of the ways in which the study of relative clauses has been biased, until quite recently, by concentration on data from English and the construction of abstract analyses to account for these characteristics. We shall not be criticizing these analyses *qua* analyses of English, but the discussion of later sections of this chapter, in which we examine relative clause equivalents in a wide range of languages, will demonstrate some of the limitations of trying to build a universal syntactic theory solely on the basis of English data and abstract analyses thereof.

One of the distinctions that has attracted considerable attention is that between restrictive and non-restrictive relative clauses in English. An example of a restrictive relative clause would be *the man that I saw yesterday left this morning*, more specifically the relative clause *that I saw yesterday* within this sentence. This clause serves to delimit the potential referents of *the man*: the speaker assumes that the sentence *the man left this morning* does not provide the hearer with sufficient information to identify the man in question (the hearer would probably have to ask *which man?*), so the additional information *that I saw yesterday* is added to indicate specifically which man is being talked about. Non-restrictive relative clauses are illustrated by the following examples: *the man, who had arrived yesterday, left this morning*, or *Fred, who had arrived yesterday, left this morning*. In these sentences, it is assumed by the speaker that the hearer can identify which man is being talked about, and that it is one particular, identifiable Fred that is being talked about, and the relative clause serves merely to give the hearer an added piece of information about an already identified entity, but not to identify that entity. In English, non-restrictive relatives (also called: appositive, descriptive, explanatory) require the relative pronoun *who* or *which*, or

their inflected forms (*whom, whose*), and are also set off intonationally from the main clause, indicated orthographically by commas. Restrictive relatives allow, in addition to *who* and *which*, the relative pronoun (or conjunction?) *that* in most instances, or even suppression of the relative pronoun/conjunction, as in *the man I saw yesterday left this morning*; moreover, it is not necessary, or usual, for the restrictive relative clause to be set off intonationally from the main clause, indicated orthographically by the absence of commas.

Despite the similar syntactic constructions for restrictive and non-restrictive relative clauses, they are radically different in semantic or pragmatic terms, in particular in that the restrictive relative clause uses presupposed information to identify the referent of a noun phrase, while the non-restrictive relative is a way of presenting new information on the basis of the assumption that the referent can already be identified. In typological terms, however, this distinction seems to be almost completely irrelevant. Formal distinction between restrictive and non-restrictive relatives is found sporadically across languages, but probably most languages have either no formal distinction, or only an intonational distinction where the relative clause follows the head noun. To give just one further example of a language with a formal distinction, in Persian the suffix *-i* is required on the head of a restrictive relative, but not on the head of a non-restrictive relative:

> *Mardhā-i* [*ke   ketābhārā be ānhā dāde budid*]   *raftand.*   (1)
> men   that books   to them you-had-given went
> 'The men that you had given the books to went.'

> *Mo'allef* [*ke   nevisandeye xubi -st*] *in*
> author   that writer   good is this
>
> *sabkrā exteyār karde ast.*   (2)
> style   has-chosen
> 'The author, who is a good writer, has chosen this style.'

In English, the semantic distinction applies equally to prenominal adjectives, as in *the industrious Japanese*, meaning (a) those Japanese that are industrious (but not those that are lazy), or (b) all Japanese, who are (incidentally) industrious; here, however, there is no formal distinction corresponding to the potential *who/that* distinction, or different intonation patterns, in the fully-fledged relative clause.

The second characteristic of work on relative clauses, especially within the generative framework (though also continuing some ideas of traditional grammar) can be seen by comparing relative clauses with the closest corres-

ponding independent sentences. For the sake of simplicity, we will for the most part give examples with the *wh-* relative forms. If we compare the relative clause *whom I saw yesterday* of *the man whom I saw yesterday left today* with the independent sentence *I saw the man yesterday*, then it is clear that there is a difference in the order of grammatical relations: in the independent sentence, as usually in English, the direct object follows the main verb; in the relative clause, however, the relative pronoun occurs clause-initially, and indeed the general principle of English relative clause formation is that the relative pronoun must occur clause-initially, or at least as part of the clause-initial noun or prepositional phrase (to allow for relative clauses like *with whom I arrived, the roof of which I repaired*). In transformational terms, relative clause formation in English involves a movement transformation, moving the *wh-* word from its normal position in the clause to clause-initial position. For present purposes, although we shall continue to use the terminology of movement, this can be understood more nuetrally as a way of referring to the difference between the word order of the independent sentence and that of the relative clause, without any necessary commitment as to the most appropriate formal means of describing this difference.

In fact, we can be even more specific about this property of English relative clauses: they involve movement without any overt trace being left behind in the position moved out of. In non-standard English, it is sometimes possible to move the *wh-* element while leaving an overt trace, in the form of a pronoun, as in *this is the road which I don't know where it leads*, compared to *I don't know where the road leads*: in the relative clause the *wh-*element *which* has been moved to clause-initial position, but the pronoun *it* has been left behind. This latter construction can be referred to as a copying transformation (movement with copying), while the kind discussed earlier is known technically as a chopping transformation (movement without an overt copy).

As described so far, English might seem to permit movement of any noun phrase to clause-initial position in the formation of relative clauses. In fact, however, there are several restrictions on this process in English. One such restriction is that it is impossible to move the subject of an embedded clause with an overt conjunction in this way, so that if we start from *I don't know where the road leads*, it is impossible to move *the road*, as subject of the embedded clause *where the road leads*, by a chopping transformation to give **this is the road which I don't know where leads*. At least, this is impossible using a chopping transformation. As we saw above, for those varieties of English where at least some relative clauses can be formed by copying, rather than chopping, relativization of such a noun phrase is possible, to give *this is the road which I don't know where it leads*. On the basis of the English data, then, one might be led to posit that

constraints on relative clause formation are constraints on movement transformations, more specifically constraints on chopping transformations. The fact that some varieties of English allow relativization of certain otherwise inaccessible positions provided a pronoun is retained would seem to reinforce this impression: the constraint is on chopping, rather than copying, so copying enables one to override the constraint.

Examination of data from a wider range of languages, however, suggests that, however adequate this may be as an analysis of relative clause formation in English, it is not adequate as a general syntactic account of relative clause formation cross-linguistically, or even of constraints on relative clause formation cross-linguistically. First, many languages quite regularly use pronoun-retention as a means of forming relative clauses (see further section 7.2.3). If constraints were simply constraints on chopping, then one would expect such languages to relativize freely any noun phrase. However, this is not the case. For instance, Zürich German requires pronoun-retention for relativization of most noun phrases, but in Zürich German it is just as impossible as in English to relativize a noun phrase which is itself within a relative clause, i.e. to start from *John saw the man that gave me the book* and relativize *the book* to give *I'm going to sell the book that John saw the man that gave me (it). On the other hand, Persian, which also has pronoun-retention, does allow the formation of such relative clauses. So constraints on chopping are not sufficient to characterize cross-language accessibility to relative clause formation.

Secondly, there are many languages in which relative clause formation does not seem to involve any syntactic movement at all, but where still there are constraints on relativization. Even for English, it is arguable that the relative clause introduced by *that* or zero involves no movement, since *that* can be analyzed as the general subordinating conjunction *that* rather than as a relative pronoun. Thus the ungrammatical relative clause *this is the road (that) I don't know where leads would involve no movement, but would still violate a constraint on relativization. The following Basque examples show that Basque has no movement in the formation of relative clauses; in these examples, -*k* indicates a (transitive) subject (ergative case) and -*ri* an indirect object (direct objects take no suffix), while the suffix -*n* on the auxiliary verb *dio* indicates that it is in a relative clause:

Gizona-k emakumea-ri liburua eman dio.   (3)
man   woman   book   has-given
'The man has given the book to the woman.'

[*emakumea-ri liburua eman dio-n*] *gizona*   (4)
'the man who has given the book to the woman'

[*gizona-k emakumea-ri eman dio-n*] *liburua*   (5)
'the book which the man has given to the woman'

[*gizona-k liburua eman dio-n*] *emakumea*   (6)
'the woman to whom the man has given the book'

Nonetheless, the range of noun phrases that can be relativized in Basque is highly restricted; in particular it is not possible to relativize on possessives, to give a literal translation of *the boy whose book the man has given to the woman*.

All of this material goes to show that, before having a reasonable chance of coming up with cross-linguistically valid generalizations about relative clause formation, it is necessary to investigate some of the different types of relative clause constructions that are found across the languages of the world. This problem is addressed in the remainder of this chapter.

7.2 TYPES OF RELATIVE CLAUSE

7.2.1 DEFINING THE NOTION RELATIVE CLAUSE

Given that the constructions which we have, so far informally, been calling relative clauses differ quite considerably in their syntactic structures across languages, it is essential that we should have some reliable way, independent of language-specific syntax, of identifying relative clauses (or at least, prototypical relative clauses) cross-linguistically. We can illustrate this by contrasting the English relative clause, discussed in section 7.1, with one type of Turkish relative clause:

[*Hasan-ın Sinan-a ver -diğ-i*]
  Hasan of Sinan to give   his
  patates-i   yedim.   (7)
  potato ACCUSATIVE I-ate
'I ate the potato that Hasan gave to Sinan.'

In terms of its syntactic structure, (7) differs considerably from its English translation. The verb form *ver-diğ-* is a non-finite form of the verb *ver* 'give', with the nominalizing suffix *-diğ*; like other nominalized verb forms in Turkish, it requires its subject (*Hasan*) in the genitive and the appropriate possessive suffix (here *-i* 'his') on the verb noun. Thus a literal translation of the head noun and relative clause *Hasanın Sinana verdiği patates* would be 'the potato of Hasan's giving to Sinan'. In

English traditional grammar, the term clause is often restricted to con-
structions with a finite verb, so in terms of this definition the Turkish
construction is not a clause, therefore not a relative clause. However, this
terminology simply reflects a general property of English syntax: subord-
ination is carried out primarily by means of finite clauses; whereas in
Turkish subordination is in general by means of non-finite constructions.
The claim found in some discussions of Turkish that Turkish does not
have relative clauses is thus in one sense correct, but from a wider per-
spective, it is clear that the Turkish construction illustrated in (7) fulfils
precisely the same function as the English relative clause: thus, in its
restrictive interpretation, there is a head noun *patates* 'potato', and the
relative clause restricts the potential reference of that head noun by telling
us which particular potato (the one that Hasan gave to Sinan) is at issue.
The lesson of this comparison is thus that we need a functional (semantic,
cognitive) definition of relative clause, on the basis of which we can then
proceed to compare relative clauses across languages, neglecting language-
specific syntactic differences in our over-all definition of relative clause,
but using them as the basis of our typology – for instance, the distinction
between finite and non-finite relative clauses is one typological parameter.

We can now be somewhat more specific about the definition of relative
clause, bearing in mind that, as so often, what we are giving is a characteri-
zation of the prototypical relative clause, rather than a set of necessary and
sufficient conditions for the identification of relative clauses. We will
assume that restrictive relative clauses are more central to the notion of
relative clause than are non-restrictives, and construct the definition ac-
cordingly. A relative clause then consists necessarily of a head and a re-
stricting clause. The head in itself has a certain potential range of referents,
but the restricting clause restricts this set by giving a proposition that must
be true of the actual referents of the over-all construction. Taking (7) as an
example, whether in its English or Turkish form, this is a relative clause
because it has a head with a range of potential referents, namely 'potato',
but the actual set of referents is limited to the potatoes (in this case, to the
one potato) of which the proposition 'Hasan gave the potato to Sinan' is
true.

In one sense, this definition is somewhat narrower than the traditional
concept of relative clause, for instance by excluding non-restrictive rela-
tives, and also certain, arguably marginal, constructions of the type *John is
no longer the man that he used to be,* where the function of the relative clause
(if such it is) is hardly to restrict the range of reference of *the man.* In
another sense, however, it is much broader. In English, for instance, it will
include not only finite relative clauses of the kind already discussed, but
also non-finite (participial) constructions like *passengers leaving on flight*

738 should proceed to the departure lounge, and even restrictive attributive adjectives, like *good* in *the good students all passed the examination*. We would argue that this is not a disadvantage: note that these latter constructions have straightforward paraphrases as relative clauses in the traditional sense (*passengers who are leaving, the students who are good*); moreover, none of the generalizations made in the discussion to follow on universals of relative clause formation is affected by either including or excluding such constructions.

One requirement on the definition of relative clause is that, in order to say that a language has relative clauses, it should be the case that there is some construction or constructions correlating highly with the definition given above. In this sense, it may well be that some languages do not have relative clauses. In Walbiri, for instance, the usual translation of ' I speared the emu that was drinking water ' would be:

Ŋaᵗʸulu-ḷu   -ṇa   yankiri pantuṇu kutʸa
I   ERGATIVE AUXILIARY emu   speared CONJUNCTION

  -ḷpa   ŋapa ŋaṇu.   (8)
  AUXILIARY water drank

However, this Walbiri sentence is equally the most natural translation of ' I speared the emu while it was drinking water ', i.e. the Walbiri could be used in answer to either 'which emu did you spear?' or 'when did you spear the emu?' In Walbiri, then, this is a fairly general subordination construction, and not a construction whose sole, or even prototypical, function is to encode meanings in accordance with our definition of relative clause. We therefore leave open the possibility that some languages do not have relative clauses, and if this possibility does have to be accepted, then such languages are irrelevant to (though not counterexamples to) generalizations made with regard to languages that do have relative clauses.

If the Walbiri construction is considered a relative clause, then it differs in one important respect from more familiar types of relative clauses: the relative clause is not a constituent of the main clause, rather it is simply attached, 'adjoined' to the main clause, whence the term adjoined relative clause; relative clauses which are constituents of the main clause are sometimes referred to as embedded relative clauses. A major typological division would thus be between embedded and adjoined relative clauses. In the typology that follows, embedded relative clauses will be the main subject of discussion, i.e. there is a primary division into embedded and adjoined relative clauses, with further subdivisions that are of relevance to the present discussion within the set of embedded relative clauses.

7.2.2 WORD ORDER AND RELATIVE CLAUSE TYPES

In chapter 4, in discussing word order typology in general, we noted the two most widespread types of relative clause with respect to word order, namely the postnominal type where the relative clause follows its head (as in English), and the prenominal type where the relative clause precedes its head (as in Turkish example (7) above). In addition to these, however, there is a third type, in which the head actually occurs inside the relative clause, and it is this type that we will illustrate in this section.

In the clearest examples of the internal-head type of relative clause, the head noun remains expressed within the relative clause, in the usual form for a noun of that grammatical relation within a clause, and there is no overt expression of the head in the main clause. The following example is from Bambara:

N ye   so   ye.   (9)
I PAST house the see.
'I saw the house.'

Tyὲ   be   [n ye   so   min ye] dyɔ.   (10)
man the PRESENT I PAST house   see build
'The man is building the house that I saw.'

In this construction, the whole clause *n ye so min ye* functions as direct object of the main clause, but the sense is clearly that of a relative clause. Bambara has SOV basic word order, therefore the main clause of (10) has the order subject – auxiliary – direct object – verb. The fact that a clause is functioning as a noun phrase referring to the head is even clearer in Diegueño, where the clause in question can take the appropriate suffix to indicate its syntactic role in the main clause, in example (13) being in the locative:

Tɘnay   ʔwa :ʔwu :w.   (11)
yesterday house I-saw
'I saw the house yesterday.'

Pwa :-pu   -Lʸ   Pciyɑwx.   (12)
house DEFINITE LOCATIVE I-will-sing
'I will sing in the house.'

[Tɘnay   Pwa :Pwu :w]-pu   -Lʸ
yesterday house I-saw   DEFINITE LOCATIVE
  Pciyawx.   (13)
  I-will-sing
'I will sing in the house that I saw yesterday.'

Here, the suffixes -*pu-L*ʸ are attached to the end of the embedded clause. This type of relative clause construction might be referred to as circum-nominal, parallel to the terms prenominal and postnominal.

In this construction, unlike most kinds of relative clause, there is no problem in processing the relative clause syntactically – it has basically the structure of a simplex sentence – but there are potential problems in working out which of the noun phrases within the relative clause is to be interpreted as its head, and therefore also its function within the main clause. In Bambara, this problem is solved by placing the relative marker *min* after the noun phrase within the relative clause that is head of that construction, as in (9)–(10). In some languages, however, there is no such marker, and relative clauses can therefore be ambiguous as to which noun phrase within them is to be interpreted as head, as in this example from Imbabura Quechua:

[*Kan kwitsa-man kwintu-ta   villa-shka*]
  you girl   to   story   ACCUSATIVE tell   NOMINALIZER
  -*ka*   *sumaj* -*mi*.   (14)
    TOPIC pretty VALIDATOR

' $\left\{ \begin{array}{l} \text{The girl to whom you told the story} \\ \text{The story that you told to the girl} \end{array} \right\}$ is pretty.'

A second kind of relative clause construction that is sometimes referred to as having an internal head is the following correlative construction from Hindi:

Ādmī ne   *jis*   *cākū se*   *murgī   ko*
man   ERGATIVE which knife with chicken ACCUSATIVE
  mārā thā, us   cākū ko   *Rām ne*   *dekhā*.   (15)
  killed   that knife ACCUSATIVE Ram ERGATIVE saw
'Ram saw the knife with which the man killed the chicken.'

The literal translation of (15) would be: 'with which knife the man killed the chicken, Ram saw that knife.' Although, in (15), the noun phrase of the first clause is repeated in the second clause, it would be possible to have a coreferential pronoun in the second clause instead of the repeated noun phrase. Clearly, this construction does have an internal head, since the relative clause *ādmī...thā* contains a full noun phrase *jis cākū* referring to the head. However, since the relative clause in such instances is not a syntactic constituent of the main clause, it is preferable to treat them, like the Walbiri type illustrated in (8) (if this is indeed a relative clause), as instances of adjoined relative clauses.

7.2.3 THE ROLE OF THE HEAD IN THE RELATIVE CLAUSE

It is clear from the definition given of relative clause in section 7.2.1 that the head of a relative clause actually plays a role in two different clauses in the over-all relative clause construction: on the one hand, it plays a role in the main clause (traditionally, the term head is often restricted to the noun phrase in question as it occurs in the main clause), but equally it plays a role in the restricting clause, i.e. the relative clause in the sense of the embedded (subordinate) clause. This is particularly clear in the correlative construction, as in (15), where an overt noun phrase appears in both clauses. More commonly, however, cross-linguistically, the head noun appears in a modified or reduced form, or is completely omitted, in one of the two clauses. The circumnomina relative clause discussed in section 7.2.2. illustrates omission of the head noun from the main clause. In this section, we shall be concerned with the expression of the role of the head noun within the embedded clause. Although, a priori, this might seem no more important than the role of the head in the main clause, it turns out that, from the viewpoint of typological variation, the encoding of the role in the embedded sentence is, cross-linguistically, one of the most significant parameters. Below, we distinguish four major types along this parameter: non-reduction, pronoun-retention, relative-pronoun, and gap.

The non-reduction type simply means that the head noun appears in full, unreduced form, in the embedded sentence, in the normal position and/or with the normal case marking for a noun phrase expressing that particular function in the clause. This type is illustrated by the Bambara (10) and Diegueño (13) examples above, i.e. by the circumnominal type.

In the pronoun-retention type, the head noun remains in the embedded sentence in pronominal form. We have already noted, in passing, that this type is found in non-standard English, as when from the sentence *I know where the road leads* one forms the relative clause *this is the road that I know where it leads*. In this construction, the pronoun *it* indicates the position relativized, i.e. enables retrieval of the information that relativization is of the subject of the indirect question clause. In English, this type has a rather marginal existence, but in many languages it is a major, in many circumstances obligatory, means of forming relative clauses, without any stylistically pejorative overtones. In Persian, for instance, pronoun-retention must be used for relativization of all grammatical relations other than subject and direct object; with direct objects, pronoun-retention is optional; with subjects, it is unusual, though examples are attested. The following examples illustrate relativization on subject, direct object, and indirect object, respectively:

> *Mard-i[ke (*u) bolandqadd bud] juje*
> man that he tall was chicken

-rā   košt.   (16)
ACCUSATIVE killed
'The man that was tall killed the chicken.'

Hasan mard-i-rā   [*ke   zan   (u -rā)*
Hasan man   ACCUSATIVE   that woman he ACCUSATIVE
zad] *mišenāsad.*   (17)
hit   knows
'Hasan knows the man that the woman hit.'

Man zan   -i-rā   [*ke   Hasan be u*
I   woman   ACCUSATIVE   that Hasan to her
sibe zamini dād] *mišenāsam.*   (18)
potato   gave I-know
'I know the woman to whom Hasan gave the potato.'

In (18), it would be impossible to omit *be u* 'to her', or *u* 'her' on its own.
    Before going on to the next type, we should note two points that emerge
from the presentation so far. The first is that a given language may have
more than one type of relative clause construction in its over-all battery of
relative clause formation possibilities. Thus (16)–(18) illustrate both gap
(see below) and pronoun-retention types within Persian, to some extent in
complementary distribution (only the gap type with subjects, only
pronoun-retention with non-direct objects), but also sometimes overlap-
ping (as with direct objects). The same can, of course, apply also to other
typological parameters, so that languages may have both finite and non-
finite types (as does English, cf. the non-finite participial construction
mentioned above), or both prenominal and postnominal types, as in Taga-
log:

babae   -ng   [nagbabasa ng diyaryo]   (19)
woman   that   reads   P newspaper

[*nagbabasa ng diyaryo-ng*] *babae*   (20)
'the woman that reads the newspaper'

The distribution of types within a language, however, is not completely
arbitrary, as we shall see in section 7.3.3. The second general point to note
is that the order of types being presented here proceeds from most explicit
to least explicit, with regard to encoding of the role of the head noun
within the relative clause. The non-reduction type is as explicit as it is
possible to be; the pronoun-retention type is less explicit, since it is

necessary to establish the appropriate anaphoric relation for the pronoun before the relative clause construction as a whole can be interpreted.

The next type is the relative-pronoun type, and is that found most frequently in European languages, although it is not particularly frequent as a type across the world's languages as a whole. As with the pronoun-retention type, there is a pronoun in the relative clause indicating the head, but instead of being in the usual position, in terms of linear word order, for a pronoun expressing that grammatical relation, it is moved to clause-initial position (occasionally preceded by, for instance, prepositions). For the pronoun in question to encode the role of the head noun within the relative clause, given that this cannot be done by order (the pronoun must be clause-initial), it is essential that the pronoun be case marked, at least to the same extent that noun phrases in main clauses are, to indicate its role. In English, those varieties of the language that distinguish nominative *who* from accusative *whom* thus count as having a relative-pronoun type of relative clause, but clearer examples can be found from languages with a richer case system, for instance Russian:

Devuška   prišla.   (21)
girl-NOMINATIVE arrived
'The girl arrived.'

devuška, [kotoraja   prišla]   (22)
girl   who-NOMINATIVE arrived
'the girl who arrived'

Ja videl devušku.   (23)
I   saw   girl-ACCUSATIVE

devuška, [kotoruju   ja videl]   (24)
girl   who-ACCUSATIVE I saw
'the girl whom I saw'

Ja dal   knigu devuške.   (25)
I   gave book girl-DATIVE
'I gave the book to the girl.'

devuška, [kotoroj   ja dal knigu]   (26)
girl   who-DATIVE I gave book
'the girl to whom I gave the book'

In each of the relative clause examples for Russian, the relative pronoun *kotor-*, though invariably in clause-initial position, unequivocally encodes the role of the head noun in the relative clause. In comparison with the

pronoun-retention type, however, this relative-pronoun type involves greater deformation of the structure of the embedded sentence: instead of appearing in the basic word order position for a subject, direct object, or non-direct object, the relative pronoun must appear clause-initially. It is for this reason that we describe this type as somewhat less explicit than the pronoun-retention type.

In some languages, unstressed pronouns, as clitic pronouns, gravitate towards sentence-second position, irrespective of their grammatical relation. This can give rise to a type of relative clause that can be classified as representing simultaneously both the pronoun-retention and the relative-pronoun types, as for instance in colloquial Czech. The normal word order is subject – verb – object:

> To   děvče uhodilo toho muže.   (27)
> that girl hit   that man
> 'That girl hit that man.'

If the object is a pronoun, it appears as a clitic immediately following the first major constituent, so that pronominalizing *toho muže* 'that man' to *ho* 'him' necessarily involves a change of word order:

> To děvče ho   uhodilo.   (28)
> that girl him hit
> 'That girl hit him.'

One way of forming relative clauses in Czech uses the invariable conjunction *co*, with a clitic pronoun referring back when relativization is of the direct or indirect object:

> muž, [co   ho   to   děvče uhodilo]   (29)
> man that him that girl   hit
> 'the man that that girl hit'

In principle, this could be regarded as pronoun-retention (the pronoun occupies the position normal for a clitic pronoun), or as relative-pronoun (the pronoun gravitates towards sentence-initial position, giving rise to a single phonological word *co-ho*, marked as accusative case). While we are not aware of any crucial typological generalizations that depend on the assignment in this particular instance, it seems reasonable to restrict the term relative-pronoun type to examples where the movement is specific to relative clauses, thus excluding examples such as the Czech example where the movement is determined by other principles.

A similar problem, which does have implications for universals of relative clause formation (see section 7.3), occurs where a language requires or allows pronouns to be present in a simplex sentence coreferential with full noun phrases, and also allows or requires the same in relative clauses. In Hausa, for instance, in simplex sentences the subject must be taken up by a coreferential pronoun:

> *Yūsufù yā zō.*   (30)
> Joseph he came
> 'Joseph came.'

*Yūsufù zō is ungrammatical. It is therefore no surprise that, when relativizing a subject, a pronoun must be retained in the relative clause:

> *dōkin [dà   ya   mutù]*   (31)
> horse that it died
> 'the horse that died'

Since the pronoun is required independently of relative clause formation, we would not classify this as an example of pronoun-retention, thereby restricting this term to examples where there is a pronoun in the relative clause in addition to possibilities of pronoun occurrence in main clauses.

Common to all of the types discussed so far has been overt indication – by a full noun phrase, pronoun, or moved relative pronoun – of the role of the head noun within the relative clause. The next major type, the gap type, simply does not provide any overt indication of the role of the head within the relative clause. In English, at least in those varieties that do not have a *who/whom* distinction, this type is used to relativize subjects and direct objects:

> *the man who/that gave the book to the girl*   (32)

> *the book which/that the man gave to the girl*   (33)

In some other languages, this type is much more widespread, and can be used to relativize a variety of even non-direct objects, as in the Korean prenominal example below:

> *[Hyənsik-i   kì kā -lìl*
> Hyensik NOMINATIVE the dog ACCUSATIVE
>   *ttäli-n]   maktäki*   (34)
>   beat RELATIVE stick
> 'the stick with which Hyensik beat the dog'

Given that this type lacks any explicit means of encoding the role of the head noun within the relative clause, the question naturally arises of how it is possible for such constructions to be interpreted reliably. In practice, a number of strategies can be invoked, ranging from strategies based on syntactic properties of the language in question to knowledge of real-world properties. In English, for instance, given that the basic, and for most purposes only, word order is subject – verb – object, a relative clause construction like *the man that saw the girl* can only be interpreted as relativizing the subject: the direct object position is already filled by *the girl*, while the subject position preceding *saw* is empty. Indeed, given the rigid word order of English, it is difficult to construct examples that are ambiguous, though it is possible to find examples such as *the model that the artist helped to paint* (cf. either (a) *the artist helped the model to paint*, or (b) *the artist helped to paint the model*). In languages with freer word order, this strategy is not available, although unequivocal interpretation will often be possible where an obligatory argument of a verb is missing. Thus in Turkish example (7), the verb of the relative clause, 'give', would normally take three arguments (subject, direct object, and indirect object); its subject and indirect object are expressed in the relative clause, so by elimination the position relativized must be the direct object. The interpretation of each of the Basque relative clauses (4)–(6) is determined likewise: this is particularly clear in differentiating between 'the man who gave the book to the woman' and 'the woman to whom the man gave the book', where the alternative interpretations 'the man to whom the woman gave the book' and 'the woman who gave the book to the man' would still make sense.

In some instances, however, neither of these structurally based strategies will work, and recourse must be had to common sense: thus in Korean example (34), the most likely relation, in the real world, between the act of someone's hitting a dog and a stick is that of instrument, so that the only natural interpretation is 'the stick with which Hyensik beat the dog', rather than 'the stick for which Hyensik beat the dog' or 'the stick behind which Hyensik beat the dog'. In yet further instances, even this strategy will fail, and the resulting relative clause will simply be ambiguous, as in the following example from Imbabura Quechua:

[*Kan shamu-shka   llajta-ka*]   *uchilla-mi.*   (35)
  you come   NOMINALIZER town   TOPIC small   VALIDATOR
'The town you are coming to/from is small.'

Since one can come either to or from a place, either interpretation is permissible. In some languages, restrictions are sometimes placed on the range of interpretations of such potentially ambiguous constructions, so

that in Korean, unlike Imbabura Quechua, example (36) can refer to the
place Hyensik came to, but not the place he came from:

    [*Hyansik -i*   *o*   *-n*]   *mikwuk*   (36)
     Hyensik NOMINATIVE come RELATIVE United-States
    'the United States, to which Hyensik came'

Such restrictions apparently have to do with less versus more marked
interpretations, but we are not aware of any detailed wide-ranging cross-
language study of such instances.

    Some further ways in which the range of interpretations can be re-
stricted within the gap type of relative clause are discussed below, with
respect to the equi type (section 7.2.4) and constraints on accessibility to
relativization (section 7.3).

7.2.4 THE ROLE OF THE HEAD IN THE MAIN CLAUSE

In most of the world's languages, it seems that the role of the head noun in
the main clause makes little or no difference to the possibility of forming
relative clauses or to the particular relative clause construction that is used.
However, there are some exceptions, which we will discuss in this section.
    One phenomenon that is found sporadically is that known to traditional
grammarians of Latin and Greek as attraction, whereby the case marking
of the head noun in one clause is attracted into that of the other clause.
Persian provides particularly clear exemplification of attraction from the
relative clause. Without attraction, we would have (37) and (38) for rela-
tivization of subject and direct object respectively:

    Ān zan   -i-rā   [*ke   diruz   āmad*] *didam.*   (37)
    that woman   ACCUSATIVE   that yesterday came I-saw
    'I saw that woman who came yesterday.'

    Zan   -i [*ke   didid*]   *injā -st.*   (38)
    woman   that you-saw   here is
    'The woman that you saw is here.'

As an alternative to (37), however, one can have (39), where *-rā* is omitted
from the direct object (though definite) of the main clause because the
head also functions as subject of the relative clause:

    Ān zan-i [*ke diruz āmad*] *didam.*   (39)

Likewise, since the head is not only subject of the main clause in (38) but also direct object of the relative clause, it can be marked with the direct object marker *-rā* in subject position too:

> *Zan-i-rā[ke didid] injā-st.*   (40)

In Ancient Greek, attraction usually works the other way round, an expected accusative relative pronoun in the relative clause being attracted into the case of its antecedent:

> *ek   tōn póleōn   [hōn   éxei]*   (41)
> from the cities-GENITIVE   which-GENITIVE he-has
> 'from the cities which he has'

The preposition *ek* requires the genitive case, so the genitive case of *tōn póleōn* is as expected in the main clause; the verb *éxei* 'he has', however, would be expected to have an accusative object, but instead the relative pronoun has been attracted into the case of the noun phrase within the main clause.

A more deep-rooted interaction between main clause and embedded clause roles is found in languages that have the equi type of relative clause, whereby the head noun must fulfil the same role in both clauses in order for a relative clause, or at least a relative clause within a certain range, to be grammatical. In some Australian languages, the general requirement that noun phrases can only be omitted if both clauses have the same subject (defined, as will be recalled from section 5.3, as P of a transitive construction or S of an intransitive construction) leads more particularly in the case of relative clause constructions with omission of the head noun in one clause to the requirement that the head be subject of both clauses. Note that since omission of the noun phrase destroys encoding of its role within that clause, the equi type is actually a subtype of the gap type of relative clause. A somewhat more widespread use of the equi type is found in Modern Hebrew: if a position that would normally require a preposition is relativized, and if moreover the same preposition also occurs on the head noun in the main clause, then the preposition (and accompanying pronoun) may be omitted, at least for many speakers, as in:

> *Natati sefer le oto   yeled [še   Miriam natna (lo)   sefer].* (42)
> I-gave book to same boy   that Mary   gave   to-him book
> 'I gave a book to the same boy that Mary gave a book to.'

Although the equi type of relative clause does occur in languages that are otherwise very different genetically, areally, and typologically, it is always a very restricted type. No language seems to have only the equi type of relative clause and also the possibility of relativizing on a wide range of syntactic positions, and it is easy to see why this should be so, from a functional point of view: in such a language it would be very easy to express such unlikely pieces of information as 'the crops prospered notwithstanding the rain notwithstanding which the excursion still went ahead', but impossible or difficult to say 'the man that you saw has already sold his house'. The equi type seems only to exist as a marginal means of shortening expressions in languages where some other type(s) is the basic relative clause type, or in languages where there are also severe restrictions on the positions that can be relativized in the relative clause. Moreover, although examples such as those of the present section do occur where the role of the head in the main clause is relevant to relative clause formation, this is always in conjunction with its role within the relative clause, i.e. the actual constraint is on the relation between main clause and embedded clause roles. Since there are many instances, discussed in sections 7.2 and 7.3, where the role of the head in the relative clause is important, this justifies the observation made at the beginning that the role of the head noun in the subordinate clause is a major typological parameter for cross-language comparison.

7.3 ACCESSIBILITY TO RELATIVE CLAUSE FORMATION

7.3.1 SIMPLEX SENTENCES

In section 7.3, we shall return to one of the problems discussed in our brief treatment of English in section 7.1, namely constraints on accessibility of noun phrase positions to relative clause formation. In the light of the discussion of section 7.2.3–4, it is clear that we will be concerned primarily, indeed exclusively, with the role of the head noun within the embedded clause. In the present section, we will restrict ourselves to relativizing constituents of simplex sentences. Here, English presents essentially no evidence of any kind of restriction, since it is possible to relativize on, for instance, subject, direct object, non-direct object, and possessor in the possessive construction as in:

the man [who bought the book for the girl]   (43)

the book [which the man bought for the girl]   (44)

the girl [for whom the man bought the book]   (45)

the boy [whose book the man bought for the girl]   (46)

In many languages, however, there are heavy restrictions on relativization on these positions. For the purposes of the present section, we will limit the discussion to precisely the four positions just mentioned, since these positions seem to form a cross-linguistically valid hierarchy with respect to relativization. Certain other positions, such as locatives and temporals, do not seem to fit into this hierarchy: in some languages they are very easy to relativize, in other languages very difficult to relativize.

The intuition that underlies the discussion of the present section is a very simple one: the hierarchy subject > direct object > non-direct object > possessor defines ease of accessibility to relative clause formation, i.e. it is, in some intuitive sense, easier to relativize subjects than it is to relativize any of the other positions, easier to relativize direct objects than possessors, etc. Clearly, however, this intuition requires much more substantiation before it can be tested out rigorously. And the particular instantiation of the intuition that seems to be borne out as a language universal is the following: if a language can form relative clauses on a given position on the hierarchy, then it can also form relative clauses on all positions higher (to the left) on the hierarchy; moreover, for each position on the hierarchy, there is some possible language that can relativize on that position and all positions to the left, but on no position to the right. In fact, in order to provide evidence for the second of these two points, we need at least one actual language to illustrate each cut-off point on the hierarchy; fortunately, however skewed the set of actual languages may be in other respects, it does provide us with the actual examples that we need.

Thus there are languages that can only relativize subjects, such as Malagasy. Starting from a simplex sentence like (47), the only possible relative clause is (48):

> *Nahita ny vehivavy ny mpianatra.*   (47)
> saw   the woman   the student
> 'The student saw the woman.'

(It will be recalled that Malagasy has verb – object – subject basic word order.)

> *ny mpianatra[izay nahita ny vehivavy]*   (48)
> the student   that saw   the woman
> 'the student who saw the woman'

Sentence (48) cannot mean 'the student whom the woman saw', nor is there any way of translating this English relative clause literally into Malagasy (we shall see below how this information can be encoded in Malagasy).

Likewise, there are languages where relativization is possible only of subjects and direct objects, such as Kinyarwanda:

> *N-a   -bonye umugabo [w   -a   -kubise abagore].*   (49)
> I   PAST see   man   RELATIVE PAST strike women
> 'I saw the man who struck the women.'

> *Nabonye abagore [Yohani yakubise].*   (50)
> I-saw   women   John   he-struck
> 'I saw the women that John struck.'

It is not, however, possible to take an instrument like *n-ikaramu* 'with the pen' in (51) and form a relative clause directly corresponding to 'the pen with which John wrote the letter':

> *Yohani yanditse ibaruwa n   -ikaramu.*   (51)
> John   wrote   letter with pen
> 'John wrote the letter with the pen.'

Continuing down the hierarchy, we find languages that allow relativization on the first three positions of the hierarchy, but not on possessors, such as the Fering dialect of North Frisian. Finally, there are languages like English which can relativize on all four positions.

In the sample of some fifty languages investigated by Keenan and Comrie in the article referenced at the end of this chapter, nearly all were in accordance with this generalization. However, there were a small number of counter-examples, in particular a number of Austronesian languages (and, even more specifically, West Indonesian languages) which, typically, allow relativization of subjects, do not allow relativization of direct objects, but then do allow relativization of non-direct objects and/or genitives, thus clearly violating the generalization given above as an absolute universal. In Malay, for instance, (52) illustrates relativization of a subject, and (53) relativization of a possessor; it is not possible to relativize on direct objects, or on (most) non-direct objects:

> *Gadis [yang duduk di atas bangku] itu kakak   Ali.*   (52)
> lady   that sit   on top bench the elder-sister Ali
> 'The lady who sat on the bench is Ali's elder sister.'

> *Orang [yang abang   -nya memukul saya] itu*   (53)
> person that elder-brother his hit   me   the
> 'the person whose elder brother hit me'

Faced with a small number of counterexamples to an otherwise valid universal, there are, as always, two ways of proceeding. On the one hand, one could simply say that the universal is a tendency, rather than an absolute: the number of exceptions is small relative to the over-all sample, moreover the fact that most of the exceptions belong to a single genetic and areal grouping serves only to accentuate their exceptional nature. The alternative would be to try and reformulate the universal, effectively weakening it, so that the counterexamples are no longer counterexamples; this is the strategy adopted by Keenan and Comrie in the work cited. They argue that, if one distinguishes different strategies of forming relative clauses, in particular if one distinguishes between (a) prenominal, postnominal, and internal-headed relative clauses, and (b) between relative clauses where the role of the head noun in the relative clause is encoded ([+ case]) versus those where it is not ([− case]), then the more general universal can be replaced by two more specific universals: (a) every language can relativize on subjects; (b) any relative clause strategy must cover a continuous segment of the accessibility hierarchy.

Given this reformulation, nearly all of the counterexamples disappear, in particular those from West Indonesian languages. One counterexample does still remain, namely Tongan, where a [+ case] strategy is used for (some) subjects and all non-direct objects and genitives, but not for direct objects, which use a [− case] strategy; however, it is possible, perhaps, to circumvent this counterexample by taking note of the relatively high degree of ergativity in Tongan, with consequent split of subject properties between A and P of the transitive construction. Let us, for the sake of argument, assume that under the reformulation there are no counterexamples. Then we have clearly succeeded in replacing a universal tendency by an absolute universal, which, other things being equal, is clearly meritorious. In the process, however, an essential part of the original intuition has been lost. The reformulation no longer corresponds to the original intuition that positions higher on the hierarchy are easier to relativize than positions lower down, since in Malay, for instance, it is clearly easier to relativize genitives than direct objects (it is in fact impossible to relativize the latter). Here we shall simply point out the advantages and disadvantages of both ways of treating counterexamples to the original generalization: there is no obvious solution to selecting which is the better, but for present purposes it is essential that the issue involved should be understood, even if its resolution is not to hand.

Given that we are using grammatical relations such as subject in stating the universal restrictions on accessibility to relative clause formation, one question that might naturally arise is whether the relevant grammatical relations are those of surface structure, or whether more abstract syntactic

analysis is required. In fact, it turns out that the relevant grammatical relations are those of surface structure: this particular piece of work thus provides evidence in favour of stating at least some universals in terms of more concrete levels of analysis. The evidence here is most easily seen in languages that have heavy constraints on relativization, such as Malagasy (only on subjects) or Kinyarwanda (only on subjects and direct objects).

In Malagasy, in addition to the active voice illustrated in (47), there are also other voices which enable other arguments of the verb to appear as surface structure subject. For instance, in the examples below, (54) is in the active; (55) is in the so-called passive, with the direct object of the active as surface subject; (56) is in the so-called circumstantial voice, with a non-direct object (here, benefactive) as surface subject:

Nividy ny vary ho an'ny ankizy ny vehivavy.   (54)
bought the rice for the children the woman
'The woman bought the rice for the children.'

Novidin'   ny vehivavy ho an' ny ankizy ny vary.   (55)
was-bought the woman for   the children the rice
'The rice was bought for the children by the woman.'

Nividianan'   ny vehivavy ny vary ny ankizy.   (56)
was-bought-for the woman   the rice the children
'The children were bought rice by the woman.'

In (54), only the subject, *ny vehivavy* ' the woman ' can be relativized. However, if one wants to convey the information of English relative clauses like ' the rice that the woman bought for the children ' or ' the children for whom the woman bought the rice ', then it is possible to do so provided one uses the appropriate non-active voice, with the relativized noun phrase in subject position. From (55) we can thus form (57), and from (56) we can form (58):

ny vary [izay novidin'ny vehivavy ho an'ny ankizy]   (57)
' the rice that was bought for the children by the woman '

ny ankizy [izay nividianan'ny vehivavy ny vary]   (58)
' the children who were bought rice by the woman '

Likewise in Kinyarwanda, in addition to sentence (51), there is an alternative construction where the semantic instrument appears as a direct object:

Yohani yandikishije ikaramu ibaruwa.   (59)
' John wrote the letter with the pen.'

In this instance, a literal translation into English is not possible; note that the change in voice is indicated by the suffix -*ish* on the verb. In this version, 'pen' can be relativized, just like any other direct object, to give a sentence with the same information content as 'I saw the pen with which John wrote the letter':

Nabonye ikaramu [*Yohani yandikishije ibaruwa*].   (60)

Note that nothing in the discussion of these Malagasy and Kinyarwanda derived sentences has required any additional principles: the behaviour of these sentences follows perfectly straightforwardly from the generalization that relativization in these languages is constrained in the way it is by surface structure grammatical relations, plus the knowledge that these languages have these voices in addition to the basic voice. Thus although it would be possible to recognize here a further type of relative clause, namely a verb-coding type (the verb form encodes what position is being relativized), it seems unnecessary to do so. One can, however, venture the further observation that heavy restrictions on relativization tend to correlate with the existence of a wide range of voices, so that positions that are not relativizable directly can be made more accessible by using a different voice. This generalization seems to be borne out, although it is a statistical correlation rather than an absolute linking of accessibility possibilities to range of voices: for instance, English has a passive voice, but this does not increase accessibility to relativization, since direct objects can be relativized directly. We may note further that in the West Indonesian languages discussed above that have a gap in relativization possibilities as defined by the hierarchy, the missing position(s) can usually be relativized indirectly by the use of a derived voice, so that the possibility of conveying this information is not lacking. In Malay, for instance, a direct object can effectively be relativized by using the passive voice, so that the noun phrase in question appears as a subject and can then be relativized:

    kawan saya [*yang dipukul   oleh Ali*]   (61)
    friend my   that hit-PASSIVE by Ali
    'my friend who was hit by Ali'

7.3.2 COMPLEX CONSTRUCTIONS

In section 7.3.1, we looked only at simplex sentences, and only at a single position at a time. However, one might venture to claim that the kinds of generalization made in section 7.3.1 can be extended if one also considers more complex constructions. In the present section, we shall examine two

instances where this does indeed seem to be the case, but also some examples where there are as yet unsolved problems.

One obvious extension would be to arguments of subordinate clauses as well as those of simplex sentences. An extension that seems to be valid is that it will never be easier to relativize a given constituent of a subordinate clause than to relativize the same constituent of a main clause, e.g. it would not be easier to relativize the direct object of a subordinate clause than the direct object of a main clause. We could reformulate this as an implicational universal: if a language can relativize a subordinate direct object, then it can relativize a main clause direct object. There are languages where one can relativize both, such as English. There are languages where one can relativize neither, such as Malagasy. And there are languages like Russian where one can relativize main clause direct objects but not subordinate clause direct objects:

devuška, [*kotoruju*   *ja ljublju*]   (62)
girl   who-ACCUSATIVE I   love
'the girl that I love'

*devuška, [*kotoruju*   *ty   dumaeš,* *čto ja ljublju*]   (63)
 girl   who-ACCUSATIVE you think   that I   love
'the girl that you think (that) I love'

In the absence of languages that can relativize subordinate direct objects but not main clause direct objects, the universal seems to hold.

A similar extension would hold with possessive constructions. In section 7.3.1, we simply asked in general whether it was possible to relativize possessors, without asking about the role of the whole noun phrase of which the possessor is a part, e.g. whether it is easier to relativize a possessor that is part of a subject noun phrase than one that is part of a direct object noun phrase. Again, there is evidence from languages in favour of this: for instance, in Malay, it is possible to relativize the possessor of a subject, but not the possessor of a non-subject:

orang [*yang abang*   -*nya memukul saya*] *itu*   (64)
person that elder-brother his hit   me   the
'the person whose elder brother hit me'

*orang [*yang saya memukul abang*   -*nya*] *itu*   (65)
 person that I   hit   elder-brother his   the
'the person whose elder brother I hit'

However, there are also problems in generalizing these extensions, indicating that further work needs to be done in some of these areas. For

instance, a reasonable extension would be to claim that subjects of subordi-
nate clauses should be more accessible to relativization than non-subjects
(just as subjects of main clauses are more accessible than non-subjects).
However, all the evidence suggests just the opposite. In English, non-
subjects of subordinate clauses are freely relativizable, while subjects can
be relativized only if there is no conjunction:

> the girl [that you think (that) I love]   (66)

> the girl [that you think (*that) loves me]   (67)

For some speakers of Hungarian, quite irrespective of conjunctions,
subordinate subjects cannot be relativized, whereas non-subjects often can:

> a pénz, [amit   mondtam, hogy a   (68)
> the money which-ACCUSATIVE I-said   that the
> fiú elvett]
> boy took-away
> 'the money that I said (that) the boy took away'

> *a fiú, [aki mondtam, hogy elvette   a pénzt]   (69)
> the boy who I-said   that took-away the money-ACCUSATIVE
> 'the boy that I said took away the money'

(*Elvette* is used with a definite direct object, otherwise *elvett*. Other
speakers of Hungarian find (69) grammatical.) In Imbabura Quechua, it is
possible to relativize a non-subject of an embedded clause using the gap
type, but not the subject of an embedded clause:

> [Marya Juan wawa -ta   riku-shka
> María Juan child ACCUSATIVE see NOMINALIZER
>   -ta   ni -shka]   llugshi-rka.   (70)
>   ACCUSATIVE say NOMINALIZER leave PAST-3SINGULAR
> 'The child that María said that Juan saw left.'

> *[Marya warmi Juan-ta   riku-shka
> María woman Juan ACCUSATIVE see NOMINALIZER
>   -ta   ni -shka]   llugshi-rka.   (71)
>   ACCUSATIVE say NOMINALIZER leave PAST-3SINGULAR
> 'The woman that María said saw Juan left.'

There is thus good cross-linguistic evidence for the surprising gener-
alization that subordinate non-subjects are easier to relativize than subor-
dinate subjects, but apparently no good explanations as to why this should
be so.

7.3.3 THE DISTRIBUTION OF RELATIVE CLAUSE TYPES

In section 7.2, we noted that it is often the case that a given language has more than one relative clause type, usually with at least some non-overlap between them. It has been observed that, in such instances, the distribution of relative clauses is not arbitrary. For instance, where the choice is between a pronoun-retention and a gap relative clause, it is nearly always the case that the pronoun-retention type is used lower down the accessibility hierarchy (or, more generally, in positions that are, cross-linguistically, less accessible), while the gap strategy is used higher up. As already mentioned, Persian uses the gap type on subjects and direct objects, the pronoun-retention type on direct objects, non-direct objects, and genitives (and also constituents of subordinate clauses and coordinate constructions). In Malay, the gap type is used for subject, the pronoun-retention type for genitives (and some non-direct objects, which are constructed basically as genitives). In English, pronoun-retention is marginal, but is used in some varieties for one of the least accessible positions, namely subject of a subordinate clause with an overt conjunction. This observation can probably be generalized somewhat: wherever a language has both a more explicit and a less explicit way of forming relative clauses (in the sense of section 7.2.3), then the more explicit type will be used lower down the hierarchy and the less explicit type higher up the hierarchy.

The generalization thus has a functional basis: the more difficult a position is to relativize, the more explicit indication is given of what position is being relativized, to facilitate recovery of this information. This is, of course, simply a generalization of the functional explanation already suggested in section 1.3.3.

NOTES AND REFERENCES

The richest account of the typology of relative clauses is Lehmann (1984). A useful source of information on relative clauses in a variety of languages is Peranteau *et al*. (1972). The Persian data are from Lambton (1957, 75-8).

The major work on constraints on movement transformations in English is Ross (1986, first distributed in 1967). The critique of stating such constraints purely as constraints on movement (chopping), with discussion of the Zürich German and Persian examples, is from Keenan (1975, 406-10). Relative clauses in Basque are discussed by de Rijk (1972).

The characterization of relative clauses proposed in section 7.2 is based on Keenan & Comrie (1977, 63-4). The Walbiri example is from Hale (1976), which also discusses similar constructions in other Australian languages.

In section 7.2.2, the Bambara examples are from Bird (1968), and the Diegueño examples from Gorbet (1976: 43–4). The Imbabura Quechua examples here and below are from Cole (1982, 47–60). The classification in section 7.2.3 follows closely that of Maxwell (1979), itself based on Givón (1975a). The Korean examples here and below are from Tagashira (1972). In section 7.2.4, the Ancient Greek example is from Goodwin (1894, 220). The Hebrew example is discussed by Cole (1976, 244).

The discussion of section 7.3.1 is a simplified and revised rendering of Keenan & Comrie (1977); examples cited or referenced in this work or its companion (Keenan & Comrie, 1979) are not referenced again here. The positions indirect object and object of comparison, for which there is less cross-linguistic evidence, are omitted here. In the original, the hierarchy constraints are presented as absolute universals (with reservations over Tongan). The Kinyarwanda examples are from Gary & Keenan (1977). My understanding of the Malay examples has benefited greatly from discussion with Chiang Kee Yeoh (Universiti Sains Malaysia, Minden, Penang), and the examples are from Yeoh (1979); the grammaticality judgements given in Keenan & Comrie (1977, 1979) are questionable, at least for the standard languages of Malaysia and Indonesia. The ungrammaticality of sentences like (67) (with *that*) has been the subject of much debate within generative grammar, starting with Perlmutter (1971, 99–122), in more recent work under the term '*that*-trace effects'; despite the amount of work on the topic, there is no generally accepted solution, as is clear from the survey in Van der Auwera (1984).