Taiwan Journal of Linguistics
Vol. 1.1, 69-96, 2003
doi:10.6519/TJL.2003.1(1).3

LINKING AND NOMINAL OMISSION

John Truscott

ABSTRACT

Most (and probably all) of the world's languages allow omission of a variety of nominal elements, ranging from solitary nouns to entire noun phrases. This omission occurs in a great many seemingly disparate contexts. The thesis of this article is that for a great assortment of such cases, in a wide variety of languages, a unified explanation can be given in terms of the notion of *linking*: Nominal omission is possible in the presence of an item that has the function of linking the nominal to an overt element in the sentence. The apparent logic of the linking-omission connection is that the link serves to point out the location of the missing item. The intimate connection between linking and nominal omission is demonstrated in a variety of constructions in a number of languages.

1. INTRODUCTION

Languages have a great many forms of nominal omission. The missing nominal can be a noun by itself, an entire noun phrase, or something in between, and the omission can occur in a variety of syntactic contexts, apparently differing widely across languages. I will argue here that despite this apparent diversity, nominal omission is actually a single phenomenon. The unifying factor, I suggest, is the presence of a *link*, i.e., an element that serves to connect the nominal to another item in the sentence. Omission is possible precisely when such an element is present.

The paper begins with a summary of linking, describing the identifying characteristics of links and the various forms in which they are realized crosslinguistically. This summary provides the background for the subsequent discussion of the omission of nouns, noun-modifier sequences, and entire noun phrases, involving a number of constructions in a variety of languages. The discussion of NP omission requires an

extension of the concept of linking to include elements outside noun phrases, which is presented at the beginning of the relevant section. Throughout, I will present the observations in a relatively theory-neutral form, as the goal is to present a phenomenon that should be accommodated in any syntactic theory, not to suggest how the accommodation should be accomplished within any given theory.

2. LINKING IN NP

In this section, I will first consider the defining characteristics of links, i.e., the criteria by which an element can be identified as a link. This will be followed by a look at the various types of elements that satisfy these criteria. (For much more detailed discussion, see Truscott 2000.)

2.1 Characterizing Links

Central to the idea of a link is the function of connecting an element within NP to the head noun of the phrase (though this is not to say that a link could not also serve additional functions—see below). Thus a link can be identified as such by its association with non-heads within the NP; in other words, a link is an element whose appearance is contingent on the presence of another non-head element.

In addition, one would expect to find a very local relation between the link and one or both of the elements it connects. The prototypical position is between modifier and noun and attached to the former. Limited deviations from this ideal will result from characteristics of individual languages. Thus, a language in which adjectives follow nouns and suffixation is favored over prefixation will naturally produce N-A-link configurations. There will also be many cases in which other material separates the modifier from the noun. In such cases, the link should be adjacent to, and probably attached to, one of the two. Typically, one would expect this to be the modifier, although the opposite should be true if the link bears features of the modifier rather than the noun.

Finally, links can be identified by their uniqueness within a given construction. It would be odd, at best, for a single modifier to be connected to the head by two distinct links (though there could well be two or more links within an NP, if that NP contains two or more

modifiers).

Links can be identified, then, by their linking function, their co-occurrence with a modifier, their proximity to one or both of the elements they connect, and their uniqueness in any given construction. Elements that have these characteristics can be found in a wide variety of languages, in somewhat different forms, depending to some extent on the type of language in which they appear.

2.2 Types of Links

2.2.1 Linking Particles

One type of link is exemplified by the Chinese particle *de*, shown in the following example.

(1) hen congming de xuesheng
very intelligent LP student

The element *de* is a *linking particle* (*LP*), defined by its function of linking the modifier *congming* to the head noun and by its invariant form; in contrast to the other link types to be considered below, its form does not depend on the identity of the noun. In addition to Sino-Tibetan languages such as Chinese, LPs are also commonly found in Austronesian (Lopez 1965; Josephs 1975; Dardjowidjojo 1978) and Indo-European (Elwell-Sutton 1941; see also the discussion below).

The English "possessive" marker –*s* is another example.

(2) Fred-'s pen

The -*s* suffix in such cases has all the characteristics of a link: Its presence is dependent on the occurrence of a modifier, which it connects to the head noun, it is unique in serving this function, and it is always adjacent to the modifier. The suffix is in fact identical to the Chinese linking particle in these respects, as well as in its lack of semantic content. In regard to this last point, the term "possessive" is misleading; as pointed out by Fries (1938) and many later authors, the suffix is consistent with a wide variety of relations between modifier and noun, one of which is possession. It is also worth noting that the Chinese linking particle is also used to connect "possessors" to head nouns in the

same way that -*s* does in English. (I will return to this Chinese construction shortly.)

2.2.2 Gender Markers and Noun-Class Markers

A second type of link is the *gender marker* (*GM*) shown in the Spanish examples below.

(3) a. la pelota roj-a
the ball red-fs
"the red ball"
b. el libro roj-o
the book red-ms
"the red book"

The feminine –*a* in (3a) and the masculine –*o* in (3b) are examples of GMs, serving to link the modifier *roj-* to the head nouns. Gender markers of this sort occur quite generally throughout the Indo-European and Afro-Asiatic (Welmers 1973) families and can also be found, less commonly, among Australian (Dixon 1980) and Caucasian (Comrie 1981) languages. In all these cases they show all the distinctive features of links.

A variety of gender marker that is sometimes treated as a distinct entity is the *noun-class marker* (*NCM*), exemplified by the following Forrest River example, taken from Capell and Coate (1984):

(4) amba a-njinga a-newur
kangaroo(s) NCM-this/these NCM-large
"this/these big kangaroo(s)"

The prefix *a-*, appearing on the determiner and the adjective, indicates the grammatical class to which the noun belongs. Nouns in Forrest River and similar languages are divided into a number of classes, loosely based on semantic criteria, in much the way that Spanish nouns are split into masculine and feminine, the main difference being that noun-class languages generally distinguish a somewhat larger number of classes. Noun-class markers are common among Australian, Caucasian (Comrie 1981) and Niger-Kordofanian languages (Welmers 1973).

In some of these languages the NCM takes a slightly different form, as

shown in the Swahili example below, taken from Welmers (1973).

(5) ki-kapu ki-kubwa ki-moja
NCM-basket NCM-large NCM-one
"one large basket"

Swahili nouns are also split into a number of classes, with the class being marked on the noun's modifiers just as it is in Forrest River. But Swahili (and other Niger-Kordofanian languages) differs from the Australian cases in that the class is also marked on the noun itself, as can be seen in (5). Given the discussion above, this phrase contains two links, namely the prefixes on the numeral and the adjective. The affix which appears on the noun is not a link, since its appearance is unrelated to the presence of any modifier in the NP. The contrast between Swahili and Forrest River can be compared to that between Spanish, which marks nouns for gender, and certain other gender-languages, such as German, in which only the modifiers show the noun's gender.

The similarity between gender markers and noun-class markers is readily apparent. It becomes even clearer when one looks at the semantic criteria by which the classes are distinguished. The North-East and North-Central Caucasian languages (Comrie 1981) provide a striking example. Each has between two and eight noun classes. In those that have only two, the distinction is consistently between human and non-human. Those that add a third class do so by dividing human into masculine and feminine, while retaining non-human as a distinct class. These are, in other words, typical gender-languages. When more than three classes occur, the masculine-feminine distinction tends to be maintained, and non-human nouns are split into various types, producing patterns typical of noun-class languages. In these families, gender-languages are clearly one type of noun-class language, that type in which the classes are masculine, feminine, and other. The masculine-feminine distinction also appears in a number of Australian NCM-languages, such as Djingili (Yallop 1982). Thus gender systems are best seen as simply one type of noun-class system, and most authors (though not all) seem to take this identity for granted (e.g., Corbett 1991).

2.2.3 Classifiers

An element related to the noun-class marker can be found in Austro-Asiatic (Jacob 1965), Austronesian (Benton 1968; Topping 1973; Sohn 1975), Sino-Tibetan (Chao 1968; Burling 1970; Haas 1942), Mesoamerican (Suarez 1983), and even Indo-European (Elwell-Sutton 1941) languages. This is the *classifier* (*Cl*), illustrated in the Chinese examples of (6).

(6) a. nei-tiao lu
that-Cl road
"that road"
b. san-zhang zhuozi
three-Cl table
"three tables"
c. ji-ge ren
several-Cl person
"several people"

Classifiers obligatorily accompany demonstratives, numerals, and certain quantifiers in Chinese. Their grammatical function is to link these modifiers to the noun, making them like the other link types considered above. They are also like gender markers and noun-class markers in that they reflect the class of the head noun, using categories that are strikingly similar to those found in noun-class languages (Allan 1977; Denny 1976; Lyons 1977).

2.2.4 Additional Characteristics of Links

The discussion to this point has classified links into three types—linking particles, gender markers (including noun-class markers), and classifiers. But links can have additional features, sometimes overlapping with these classifications. One very common example, cross-linguistically, is number. Even English, which is somewhat impoverished in terms of linking, uses number in this way:

(7) a. this/that book
b. these/those books

The number marking clearly serves to link the demonstratives to the head noun in these cases. It has all the characteristics of a link and

therefore should be seen as one.

The person feature can also be found in some cases, when the link's features are determined by a nominal modifier rather than by the head noun. An example from Palauan (Georgopoulos 1985) is shown below.

(8) ng-meringel [$_{NP}$ a im-al a Droteo]
3s-hurt nom hand-3s nom Droteo
"Droteo's hand hurts."

Here the possessor, *Droteo*, is linked to the head by means of an agreement marker on the noun. This marker shows the person and number of the modifier. It has all the characteristics of a link and therefore should be seen as one.

2.2.5 The Unity of Linking Phenomena

Truscott (2000) argued in detail that the various types of links are all essentially the same element, the superficial differences among them following from extraneous factors, especially the morphological type of the languages in which they appear. The argument was based on the striking similarities among them, their complementary distributions, and the way the categories blend into one another, as well as the fact that all the apparent contrasts can be readily explained in terms of extraneous factors.

Consider two examples, involving linking particles. LPs appear to differ qualitatively from the other link types in that they have no semantic content; in other words, they do not distinguish classes of nouns. But this difference is exactly what should be expected if they actually are the same element. The number of classes distinguished by the other types varies widely, with the highest numbering in the hundreds and the lowest being two. The logical minimum, however, is not two, but one: we can well imagine a linking system which places all of the language's nouns in a single class. In such a system there would be only one, invariable particle, which would otherwise look exactly like the other link types. Since this precisely describes linking particles like Chinese *de*, it is reasonable to conclude that these particles are simply one type of link, that type in which the number of classes reaches its minimum value of one.

As an example of the complementary distributions of links, consider

the case of Persian. This language has both a linking particle and a very simple system of classifiers. The former is used quite generally to connect modifiers to head nouns, while the latter appears specifically in numeral constructions, as in the following example, taken from Elwell-Sutton (1941):

(9) char ta bil
four Cl spade
"four spades"

The important point is that the linking particle cannot be inserted in this construction, either before or after the classifier. This is, in a sense, a surprising fact, since the use of the particle is in general quite free. It is particularly surprising in that the Cl *ta* is also a noun, meaning "unit". The Persian linking particle is normally used in N-N constructions; thus the fact that it cannot be inserted after *ta* in (9) calls for some explanation. If both classifiers and linking particles are examples of links then the explanation is straightforward: the structure already contains a link, so a second link cannot be added.

3. OMISSION OF N

The notion of linking just described provides the key to an understanding of nominal omission. In this section I will examine omission of solitary nouns in a variety of languages and constructions, showing that it is dependent on the presence of a link to the dropped noun. I will consider cases involving linking particles, gender/noun-class markers, classifiers, and number agreement.

3.1 Linking Particles

The connection between links and nominal omission can be seen, first of all, in the use of LPs, as in the Chinese example below:

(10) a. hongse-de maozi
red-LP hat
b. hongse-de
"red one"

(11) a. hongse maozi
red hat
b. *hongse

Omission of the head noun, *maozi* ("hat"), is possible in the presence of the link, *de*, as shown in (10). It is not possible when the LP is absent, as in (11). This contrast is clearly not a matter of whether the omitted noun can be identified, or of whether its features can be recovered (see below), as the LP carries no information whatsoever about its identity. It simply serves to point out the location of the missing noun.

I noted above that the English "possessive" marker, -*s*, is a linking particle. It should follow that a possessed N can be omitted in English. This is correct.

(12) A: I need a pen.
B: Why don't you borrow Fred's *e*?

The possessed N, linked to the possessor by an overt link, can and perhaps must be dropped in this and all similar sentences.

Turkish (see Underhill 1976) provides some interesting data relating to this type of omission. Turkish adjectives do not normally bear any link to the noun they modify; as a result, nouns cannot simply be omitted after adjectives. But if a possessive suffix (marked third person) is added to the adjective, it then becomes possible to drop the noun. Thus the possessive marker is clearly licensing the noun's omission.

In this context, consider the possessive pronouns of English. In general, they differ from full NP possessors in that they do not take the link -*s*.

(13) my book
your book
his book
her book
its book
our book
your book
their book

Not surprisingly, then, omission of the following noun is not possible. The addition of the possessive link makes all the difference in the world. The possessor can now be used without the possessed noun being explicitly present, as with Turkish adjectives.

(14) mine
yours
his
hers
its
ours
yours
theirs

Despite the existence of the suppletive form *mine* (= "my one"?), the pattern is clear: A following noun can be omitted if and only if the pronoun bears the possessive affix. Thus English, like Turkish, appears to have a special process which makes omission possible by adding an overt link where needed.

This view of possessive pronouns also provides an explanation for the contrast shown below:

(15) a. Whose book is it?
b. Whose *e* is it?

(16) a. What book is it?
b. *What *e* is it?

If possessive pronouns like *hers* are to be analyzed as *her* + *s*, then *whose* should naturally be analyzed in the same manner, as *who* + *s*. Thus there is an overt link in (15), explaining the fact that omission is possible. In contrast, *what* does not show any link; as a result, omission is not possible in (16).[1]

---

[1] The behavior of *which* seems puzzling, at least on the surface. It does not show any link, yet omission is apparently possible.

(i) Which *e* do you want?

This peculiarity is accompanied by another—the fact that *one* can optionally follow *which*.

The facts are similar for Chinese. Here there are no distinctive forms for possessive pronouns. Instead, personal pronouns are used, normally with the link *de*. In such constructions it is possible to drop the possessed noun, as expected.

(17) a. wo-de maozi
I-Link hat
"my hat"
b. wo-de *e*
I-Link
"mine"

There are certain cases in which the link can be omitted, but when this is done ellipsis becomes impossible.

(18) a. wo muqin
I mother
"my mother"
b. *wo *e*
my

Again, this is as expected, since a link is present in (17), but not in (18).
The link *de* is also used when the possessor is a lexical NP, as in (19).

(19) Zhangsan-de shu

(ii) Which one(s) do you want?

In general, the use of *one* where omission can occur is at best awkward.

A likely explanation is that (i) actually does not involve omission; instead, *which* in this case is an interrogative pronominal, comparable to *who(m)*. The existence of such an element is strongly suggested by sentences of the following type:

(iii) Which would you rather have for dessert, pudding or a piece of cake?

Treating this as ellipsis (Which dessert?/food?/item? would you rather have for dessert?) seems awkward at best. Thus, *which* appears to be a genuine pronominal in such cases rather than a determiner followed by a null nominal. The same analysis is feasible for (i). The appearance that it involves a missing nominal comes from the inherent meaning of *which* (the pronominal and the determiner), necessarily involving a selection from salient items in the context and therefore precluding the sort of generic (non-ellipsis) reading sometimes found with words such as *what* and *that*. If this analysis is right, the facts for determiner *which* are exactly as predicted: Because it does not bear a link, it does not allow omission but rather requires the use of *one*, as in (ii).

Zhangsan-LP book
"Zhangsan's book"

Not surprisingly, the possessed noun can be dropped in such phrases.

(20) Zhangsan-de *e*
Zhangsan's

3.2 Gender/Noun-Class Markers

The same phenomenon can be seen with gender agreement, as illustrated by the Spanish example below:

(21) a. la pelota roj-a
the ball red-fs
"the red ball"
b. la *e* roj-a
the red-fs
"the red one"

The adjective agrees with its noun in gender (as well as number). As a result the noun, *pelota*, can be omitted in this phrase, and in all others like it. This situation occurs quite generally throughout the languages of Europe—most of these languages have A-N agreement and allow nouns to be dropped when an adjective is present.

English differs from Spanish and similar languages in that it lacks A-N agreement.

(22) the red ball/balls

The form of the adjective does not reflect the gender (or number) of the noun. As expected, English does not allow omission of the noun in such cases.

(23) a. *the red *e*
b. the red one

If the noun is omitted then the addition of the pronominal *one* is necessary, as in (23b). In contrast, gender languages such as Spanish have no element comparable to English *one*. This lack is expected, since

the analog of (23a) is grammatical in these languages and there is therefore no need for any such element.

It is important to distinguish between these cases of nominal omission and a similar but clearly distinct phenomenon. Consider the following cases:

(24) a. The old are society's greatest asset.
b. Only the good die young.

The two phrases, *the old* and *the good*, resemble the unacceptable phrase *the red* of (23a) and the acceptable *la roja* of (21b). They are fundamentally different, though, in that the missing N in (24) can only be read as "people," while that in (21b) takes whatever reference the context provides for it. The existence of a set meaning for *the old* and *the good* suggests that a lexical process is involved in these cases. In other words, they are frozen expressions, listed in the lexicon. In contrast, the absence of a fixed meaning in the previous cases suggests a syntactic derivation for (21b). It is the syntactic phenomenon in which I am interested.

Returning to (23a), I have attributed the ungrammaticality of the phrase to the lack of any inflection on the adjective. This view is based on the contrast between English and languages like Spanish, which have the relevant inflection and allow nouns to be omitted. Its validity can also be shown by a quick look at the history of English. (The following description is based on Strang 1970.) In Old English, adjectives were inflected to agree with the noun they modified, so we would expect omission of the noun to be possible in that period. The prediction is correct: In Old English, phrases like (23a) were indeed acceptable. During the Middle English period, the inflections gradually disappeared. At the same time, phrases like (23b), involving the use of *one*, gradually replaced those like (23a), clearly showing the connection between adjectival agreement and omission of nouns.

At this point the apparent conclusion is that a noun can be omitted when its modifier contains information that permits the recovery of the noun. If this is the correct characterization, then "recovery" must be used in a very weak sense, as the information borne by the adjective is far from sufficient to identify the missing noun. In (21b), for instance, fully half the singular nouns of Spanish could qualify as the missing element in the phrase. Together with the facts noted above regarding omission

after LPs, this observation suggests that identification of the following noun is not a requirement for nominal omission. What is necessary, rather, is that there be some overt link with the missing noun, regardless of the nature of this link.

Returning to possessive pronouns, in Spanish they do not bear a possessive marker, but do show agreement with the possessed noun in gender and number. Since these constitute an overt link, omission should be permitted. This prediction is correct.

(25) a. la casa mia
the house my(fs)
"my house"
b. la *e* mia
the my(fs)

Thus it appears that omission can occur whenever the noun is linked to the possessor.

Some striking support for this view of possessive pronouns comes from German. German morphologically distinguishes three genders, two numbers and four Cases, making a total of 24 inflections. Of these, three are null (the masculine singular nominative and the neuter singular nominative and accusative). When any of the 21 overt inflections is present omission can occur, with no changes in the form of the pronoun.

(26) a. Das ist mein-e Seife.
That is my-fsnom soap.
b. Das ist mein-e *e*.

In these sentences *mein-e* is marked feminine singular nominative, and omission can freely apply.

In the case of the three null affixes, however, the accompanying noun cannot simply be omitted, as illustrated by (27).

(27) a. Das ist mein Hut.
That is my hat.
b. *Das ist mein *e*.
c. Das ist mein-er *e*.

In order for the sentence to be acceptable, an overt inflection (one

appropriate for that particular gender-number-Case combination) must be added to the modifier, as in (27c). This added inflection is not simply a disambiguating device, since a great deal of ambiguity also appears among the 21 overt inflections, none of which require any adjustments when omission occurs. The affix *-e*, for instance, represents seven distinct gender-number-Case combinations. Thus *mein-e* in (26b) is seven ways ambiguous, but this ambiguity does not affect the possibility of omission.

German, then, is like Turkish and English in having a special process that serves to license nominal omission. The facts regarding possessive pronouns in these languages provide strong evidence that it is the presence of an overt link that makes the omission possible.

As noted above, noun-class markers are inseparable from gender markers, so little additional discussion is needed. I will simply note that omission of a head noun is acceptable in the presence of an NCM on an accompanying modifier, and that the phenomenon is quite general, across NCM languages.

3.3 Classifiers

Omission is also possible in the presence of the fourth type of link, the classifier. In the Chinese examples of (28), for instance, the nouns are optional, assuming only that the context is sufficiently clear.

(28) a. san-zhang zhuozi
three-Cl tables
b. ji-ge ren
how many-Cl people
c. na-ben shu
that-Cl book

And Chinese is by no means exceptional in this regard. Every classifier language of which I am aware allows free omission of the noun when it is preceded by one of these links.

3.4 The Use of Number as a Link

The number feature also serves as a link between nouns and modifiers, as in the contrast between *this book* and *these books*, noted above. When

it is present, omission of the noun is allowed.

(29) I'll take these books and you take these/those *e*.
The demonstrative is marked with one of the noun's features, plural. This overt feature is a link to the noun; thus the licensing requirement for omission is satisfied.

If this analysis is correct then the removal of the agreement feature should make omission impossible. In other words, if the inflected demonstratives in (29) are replaced by their uninflected stem forms then the sentence should become ungrammatical. Given the natural assumption that *this* and *that* are the stem forms, this prediction is confirmed.

(30) *I'll take this/that book and you take this/that *e*.

The sentence is good only if *this/that* is given a generic meaning, "this/that thing." With the meaning that is relevant here, "this/that book," it is not. This result is expected if nominal omission depends on the presence of a link, which is available in (29) but not in (30).

If this explanation is correct, then the singular-plural contrast of (29) and (30) should not hold in a language that marks demonstratives for gender, because the gender marking is itself a link, and therefore obviates the need for a plural link. The prediction is correct for Spanish.

(31) No quiero esta pelota, quiero esa *e*.
not want this(fs) ball, want that(fs)
"I don't want this ball, I want that one."

The singular demonstrative *esa* is marked feminine. As a result, the noun can be dropped, even though *esa* is singular, providing additional evidence for the analysis of English demonstratives given above.

Thus the notion of linking provides the key to an understanding of noun omission; it is dependent on the presence of an overt link. The connection holds for gender/noun-class markers, linking particles, classifiers, and number agreement. In each case, there is a nearly perfect association between the presence of an overt link and the possibility of omitting the noun.

4. OMISSION OF MODIFIER-NOUN SEQUENCES

In all the examples to this point, it has been a solitary noun which is omitted. But other elements in the NP can be dropped along with the noun, as in the following example.

(32) a. I don't want these red pens; I want those *e*.
b. I forgot my red pen, so I'll use Fred's *e*.

In each case, *e* can refer to *red pens*, rather than just *pens*; in other words, modifier + noun sequences can be dropped.

The same is true for Spanish:

(33) No quiero estos libros rojos; quiero esos *e*.
not want these books red want those
"I don't want these red books; I want those."

The empty category can be interpreted as *libros rojos* ("red books"), again indicating that the omitted element in such cases need not be simply N.

Similar facts can be found in Chinese.

(34) Wo bu yao neige hongse (de) maozi; yao zheige *e*.
I not want that red (link) hat; want this
"I don't want that red hat; I want this one."

As in the Spanish example, *zheige e* ("this e") can be used to mean "this red hat," indicating once again that it is not just the N that can be dropped. The phenomenon is quite general across languages and link types; the omission can include a modifier in addition to the noun itself. Most importantly, in all such cases the omission is licensed by the presence of a link with an accompanying modifier.

Consider, finally, the case of multiple links within a noun phrase, and therefore multiple possibilities for omission within that phrase. In the Spanish case above,

(35) estos libros rojos
these books red

the GM on *estos* connects it to *libros rojos* and therefore allows omission of the latter, as in example (33). At the same time, the GM on *rojos* connects it to *libros*, licensing its omission:

(36) estos *e* rojos
these red

The two distinct omissions are made possible by the presence of two distinct linking relations within the phrase. The same analysis applies to the Chinese example (34), the two links being the Cl *ge* and the LP *de*, each of which licenses omission of the following element(s).

But for the English phrases,

(37) a. these red pens
b. Fred's red pens

the situation is quite different. There is only one link in each phrase, the plural marker on *these* in (37a) and the possessive marker on *Fred* in (37b). No link connects *red* to *pens* in either case. Thus the unacceptability of the following phrases is expected:

(38) a. *these red *e*
b. *Fred's red *e*

5. OMISSION OF NP (PRO-DROP)

5.1 Linking Across Categories

There is a great deal of evidence that linking is a general phenomenon, not confined to NP. It includes all the link types, in a large variety of constructions, but I will limit the discussion here to examples that are relevant to omission of NPs, namely cases of agreement between a verb and its arguments. (For additional discussion of links outside NP, see Truscott, 2000.) Verbs in a variety of North American languages (Carter 1976; Haas 1948; Hoijer 1945; Krauss 1968), for instance, contain classifiers (or perhaps noun-class markers—the distinction is difficult to draw) indicating the class of their subject. In Bantu languages the noun-class markers that act as links within NP also appear on verbs, as in the

following Swahili example, taken from Welmers (1973):

(39) ki-kapu ki-kubwa ki-moja ki-lianguka.
NCM-basket NCM-large NCM-one NCM-fell
"One large basket fell."

This type of agreement is also common among Caucasian languages (Comrie 1981). Furthermore, gender markers (if they are to be distinguished from noun-class markers) can be found on verbs in Punjabi (Ganathe 1986), Hindi (Srinivasachari 1988), and Lusatian (de Bray 1980).

Thus agreement between subject and verb is strikingly similar, cross-linguistically, to agreement between noun and modifier, suggesting that agreement markings on verbs should be seen as links. This conclusion is supported by several additional similarities between nominal links and verbal agreement. First, they are parallel in their functions, agreement serving to connect subjects and verbs in the same way that links connect modifiers to nouns. Another common feature is that each is commonly associated with a particular Case, nominative for verbal agreement and genitive for links. This Case is assigned to any NP with which the element appears. Furthermore, both are normally bound, though each can occasionally be found free. In addition, each often includes the number feature. Finally, each is non-overt in many languages. These similarities point to the conclusion that verbal agreement is a link appearing outside of NP.

The situation is similar for agreement between verbs and their objects. Many languages show agreement of this sort in essentially the same way that they (and other languages) show subject-verb agreement. The natural conclusion is that there is a link connecting the verb to its object. This object link, like its subject counterpart, is normally bound but occasionally appears as an independent element, as is sometimes the case in Kiyaka (Kidima 1987), for instance.

It can be concluded, then, that links are a general feature of XP. In the following section I will explore the implications of this conclusion for omission of noun phrases.

5.2 Omission of NP

Omission of entire noun phrases has been extensively studied, under the heading of "pro-drop" or "null subjects". The phenomenon is

illustrated by the following Spanish example:

(40) *e* habla con Pancho.
talk-3s with Pancho
"He/she/it is talking with Pancho."

In Spanish and many other languages, subject pronouns can be omitted. In addition, a number of languages allow similar omission of pronouns in object position, as in the following example from the Australian language Western Desert (Dixon 1980):

(41) *e* pu + ngku + rna + nta *e*
hit FUT 1sSUBJ 2sOBJ
"I will hit you."

In such sentences, both pronouns are normally omitted. When they are included it is for the purpose of emphasis, as is the case for subject pronouns in languages like Spanish.

Less frequently discussed are cases in which objects of prepositions are omitted. The phenomenon occurs in Irish, as shown by McCloskey and Hale (1984) and McCloskey (1986), and in Navajo, as described by Platero (1982). Another type of pro-drop not involving verbal agreement can be found in Palauan, as described by Georgopoulos (1985), and Turkish (Underhill 1976), as well as in Irish (McCloskey and Hale 1984; McCloskey 1986), Navajo (Platero 1982), and a few additional languages described by Abney (1987). Possessed nouns in each of these languages agree with their possessor in person and number, as shown in the Palauan example (8) above, repeated here as (42).

(42) ng-meringel [$_{NP}$ a im-al a Droteo]
3s-hurt nom hand-3s nom Droteo
"Droteo's hand hurts."

The possessed noun is marked third person singular, agreeing with the possessor, *Droteo*. When the possessor is a pronominal, as in the following example (based on Georgopoulos' (3a)), pro-drop occurs.

(43) [$_{NP}$ a ngok-el *e*] a ?elam
[nom flute 3s] broken
"His/her/its flute is broken in two."

Georgopoulos' summary of these and related facts is that pro-drop can occur in any position which is governed by a lexical category which agrees with the omitted pronoun. Similar observations can be made about Chamorro, as described by Chung (1984). Pronouns acting as subjects, possessors, and agents of passive sentences can all be dropped. As usual, the omission occurs in the presence of agreement in each case.

Traditional explanations for pro-drop are based on its close association with "rich" verbal inflection, providing overt information about the subject. Thus in (40) and (41) the verb is marked for the person and number of its subject. This, it appears, is what makes omission of the pronoun possible. Similar observations have been made in regard to the omission of object pronouns. Omission is closely associated with the existence of agreement between the verb and its objects. Thus the agreement marker *-nta* in (41) appears to license the omission of the object, in the same way that the subject agreement marker *-rna* licenses omission of the subject. Note that English, with very little subject-verb agreement and no verb-object agreement, does not allow either type of pro-drop.

Perhaps the most striking demonstration of the connection between agreement and pro-drop comes from Pashto. Huang (1984) observed that in Pashto verbs sometimes agree with their subject and sometimes with their object, and that this agreement correlates precisely with the possibility of dropping pronouns—exactly those arguments with which the verb agrees can be omitted, regardless of whether they are subjects or objects.

Because of this association with agreement, theories of pro-drop have traditionally been based on the idea that the contents of the dropped element must be identified by verbal inflection, or must be recoverable on the basis of this inflection (e.g., Adams 1987; Huang 1984; Jaeggli and Safir 1989; Platzack 1987; Rizzi 1982). This standard view is successful in capturing the fact that there is a connection between pro-drop and agreement, though of course it must be extended to include cases in which a possessor or the object of a preposition is omitted (see above).

More importantly, though, the claim that the dropped element must be recoverable, or that its features must be identified, is rather misleading. In the typical case, verbal inflection permits only partial recovery, or identification. In Spanish, for example, person and number can be

recovered, but gender cannot. Cases occur in Chamorro (Chung 1984) in which number alone licenses pro-drop. Platzack (1987) described a Swedish dialect in which the same is true. He also showed that Faroese, which has number agreement, but only very minimal person agreement, is a pro-drop language. Furthermore, in some Russian constructions gender and number can be recovered, but person cannot. In Hungarian, object pro-drop is licensed by verbal agreement that indicates whether a third person definite object is present (Kiss 1987); in other words, the person of the object can be identified only partially and its number cannot be recovered at all.

Thus pro-drop does not require full recoverability, nor is there any particular feature or set of features that must be identified. So the claim that an omitted element must be recoverable, or that its features must be identified, presumably does not mean that it must be *fully* recoverable or that *all* of its features must be identified, nor does it mean that any particular feature(s) must be identified. What it does mean, though, is not entirely clear.

It is clear, though, that there is a connection between pro-drop and the presence of agreement features on the verb. Thus the identification idea cannot simply be dismissed. The apparent conclusion, then, is that a weaker version of identification is required. The weaker version, I suggest, is precisely the principle that underlies the omission of nouns and modifier-noun sequences: nominal elements, including full NPs, can be omitted in the presence of a link. This principle can straightforwardly accommodate all the standard cases of pro-drop, as the agreement elements are links.

6. SOME OPEN QUESTIONS: THE ROLE OF FORMAL THEORY

I have discussed nominal omission in a relatively atheoretical way in order to demonstrate the existence of a significant generalization, one which should be accommodated in any theory, and to provide a point of departure for a more formal account which would tie linking more closely to existing linguistic theories. One value of developing a formal account is that it creates the possibility of accommodating phenomena that appear relevant but on the surface at least do not fit well with the proposed generalization. In this section I will consider two sets of such cases and suggest that these sorts of issues are best approached through

the development of more formal accounts.

6.1 Pro-Drop Without Links and Links Without Pro-Drop

The principle that omitted nominals must be pointed out by links accounts for all the traditional cases of pro-drop. In two respects, though, it does not provide a complete account of the phenomenon. First, it has long been recognized in the pro-drop literature that a rich agreement system is not strictly necessary for omission. A number of languages—Chinese, for example—have no verbal agreement but still allow omission of subject and object pronouns. Linking obviously does not provide an explanation for this phenomenon. But the same limit applies, a fortiori, to identification principles; this limit therefore does not pose any problems peculiar to the linking account of pro-drop. Similarly, it has long been recognized that rich agreement systems—and therefore links—exist in some languages that do not allow null subjects; German is an example. Parallel cases can be found for omission of nouns and noun-modifier sequences; in Tagalog, for instance, the linking particle does not seem to license omission of the accompanying noun.

The implication, for both types of cases, is that something more must be said, beyond the core observation that linking is the crucial factor in nominal omission. Clearly, other factors interact with this principle in ways that are not readily apparent on the surface. Answers for troublesome cases of this sort must therefore be sought in the continuing development of formal theories of the phenomena.

6.2 Post-Nominal Possessors

The second example of the need for formal theory is not so much an apparent counterexample as a set of complex phenomena that appear to interact with linking and omission in confusing ways. English possessors sometimes appear to the right of the noun, preceded by *of*, as in the following example.

(44) a friend of Albert Einstein

The *of* looks a great deal like a link, so one might expect omission of the following NP, *Albert Einstein*, to be acceptable, contrary to fact.

The immediate explanation is that *of* is not in fact a link, as it does not

meet one of the requirements described above—uniqueness. An acceptable variant of (44) is the following,

(45) a friend of Albert Einstein's

in which the possessive *–s*, identified above as a link, co-occurs with *of*. So *of* should not be considered a link and the impossibility of omitting *Albert Einstein('s)* is not a counterexample to the generalization that links license omission.

But the construction raises a number of rather complex issues, nonetheless. First, the link *–s* does not license omission of *friend*, as it does in ordinary possessive constructions. This limit might be attributed to a (possibly language-specific) constraint on the direction of licensing. All the English omission cases considered here have involved nominals that *follow* the link. Alternatively, perhaps *of* cliticizes onto *friend* and omission of *friend* would therefore require omission of *of* as well. If the latter assigns Case to a following NP, as is commonly believed, omission might then prevent Case assignment, explaining its unacceptability. This possibility is tied up with the question of the relative points in the derivation at which Case requirements and licensing requirements must be satisfied.

Light could probably be shed on these and other issues through deeper analysis of the post-nominal possessive construction, the nature of linking, and perhaps the nature of the omitted element. These possessives are clearly marked constructions, but it is far from clear what exactly their structure is and how characteristics of this structure might interact with a more formal account of link-based omission. The precise formal character of links is similarly unclear. One might directly postulate a category *Link*, perhaps comparable to familiar functional categories. Alternatively, the set of links could be seen in a more abstract way as a cluster of functional categories with overlapping functions. If links do represent a category in themselves, this category might be either a head, having its own projection system, or simply a single item appearing between a head and each peripheral element in its phrase. The answers to these questions, which are far beyond the scope of this study, will presumably have substantial implications for the treatment of omission, both in the post-nominal possessive case and beyond, and perhaps for a range of other theoretical issues as well.

7. CONCLUSION: UNITY

I have considered omission of lone nouns, modifier-noun sequences within NP, and full NPs. Each is dependent upon the presence of a link connecting the omitted item to an overt element in the sentence, in effect pointing out the location of the missing item. As the phenomena all involve omission of the same type of elements (nominals) under the same conditions, the natural conclusion is that they are a single phenomenon.

The idea that the location of a dropped nominal must be pointed out by something, namely a link, is a weaker and at the same time more general version of the standard identification idea in the pro-drop literature. Instead of making the overly strong requirement that the nominal's features be recoverable, it states simply that the location of a dropped nominal must be overtly indicated. There is a certain redundancy in this function, since the existence of a gap where a nominal should be would also serve to point out the missing element. It is a natural and non-problematic sort of redundancy, though, since it is the sort that facilitates processing. Languages are, in fact, full of such redundancy, as should be expected in view of the fact that speech is an imperfect means of transmitting information and that some of the information is often lost as a result.

Finally, it should be emphasized that this discussion has provided only a preliminary understanding of the linking-omission connection. In the preceding section I suggested that there is a need for more formal explanations and that apparent counterexamples remain. Problem cases of this sort are interesting and can provide directions for future research, particularly involving more formal accounts of the phenomena. But they should not be allowed to obscure the fact that a great wealth of evidence points to an intimate connection between linking and nominal omission, and therefore to the unity of nominal omission phenomena.

REFERENCES

Abney, Steven P. 1987. *The English NP in its Sentential Aspect*. Cambridge: MIT dissertation.

Adams, Marianne. 1987. From Old French to the theory of pro-drop. *Natural Language*

and Linguistic Theory 5: 1-32.

Allan, Keith. 1977. Classifiers. *Language* 53: 285-311.

Benton, Richard A. 1968. Numeral and attributive classifiers in Trukese. *Oceanic Linguistics*, 7: 104-146.

Burling, Robbins. 1970. *Man's Many Voices: Language in its Cultural Context.* New York: Holt, Rinehart, and Winston.

Capell, A., and Coate, H. H. J. 1984. *Comparative Studies in Northern Kimberley Languages.* Department of Linguistics, Research School of Pacific Studies, Australian National University.

Carter, Robin M. 1976. Chipewyan classificatory verbs. *International Journal of American Linguistics* 42: 24-30.

Chao, Yuen Ren. 1968. *A Grammar of Spoken Chinese.* Berkeley: University of California Press.

Chung, Sandra. 1984. Identifiability and null objects in Chamorro. *Proceedings of the Tenth Annual Meeting of the Berkeley Linguistic Society*, ed. by Claudia Brugman and Monica Macaulay, 116-130. Berkeley, CA: Berkeley Linguistic Society.

Comrie, Bernard. 1981. *The Languages of the Soviet Union.* Cambridge: Cambridge University Press.

Corbett, Greville G. 1991. *Gender*. Cambridge: Cambridge University Press.

Dardjowidjojo, Soenjono. 1978. *Sentence Patterns of Indonesian.* Honolulu: University of Hawaii Press.

de Bray, R. G. A. 1980. *Guide to the West Slavonic Languages.* Columbus, OH: Slavica.

Denny, J. Peter. 1976. What are noun classifiers good for? *Papers from the Twelfth Regional Meeting of the Chicago Linguistic Society*, ed. by Salikoko S. Mufwene, Carol A. Walker, and Sanford B. Steever, 122-132. Chicago: Chicago Linguistic Society.

Dixon, R. M. W. 1980. *The Languages of Australia.* Cambridge: Cambridge University Press.

Elwell-Sutton, L. P. 1941. *Colloquial Persian.* London: Routledge and Kegan Paul.

Fries, Charles C. 1938. Some notes on the inflected genitive in present-day English. *Language* 14: 121-133.

Ganathe, N. S. R. 1986. *Learn Punjabi in 30 Days* (3rd ed.). Madras: Balaji.

Georgopoulos, Carol. 1985. Variables in Palauan syntax. *Natural Language and Linguistic Theory* 3: 59-94.

Haas, Mary R. 1942. The use of numeral classifiers in Thai. *Language* 18: 201-205.

Haas, Mary R. 1948. Classificatory verbs in Muskogee. *International Journal of American Linguistics* 14: 244-246.

Hoijer, Harry. 1945. Classificatory verb stems in the Apachean languages. *International Journal of American Linguistics* 11: 13-23.

Huang, C.-T. James. 1984. On the distribution and reference of empty pronouns. *Linguistic Inquiry* 15: 531-574.

Jacob, Judith M. 1965. Notes on the numerals and numeral coefficients in Old, Middle and Modern Khmer. *Lingua* 15: 143-162.

Jaeggli, Osvaldo, and Safir, Kenneth J. (Eds.) 1989. *The Null Subject Parameter.* Dordrecht: Kluwer.

Josephs, Lewis S. 1975. *Palauan Reference Grammar.* Honolulu: The University Press of Hawaii.

Kidima, Lukowa. 1987. Object agreement and topicality hierarchies in Kiyaka. *Studies in African Linguistics* 18: 175-209.

Kiss, Katalin E. 1987. *Configurationality in Hungarian.* Dordrecht: Reidel.

Krauss, Michael E. 1968. Noun-classification systems in Athapaskan, Eyak, Tlingit, and Haida verbs. *International Journal of American Linguistics* 34: 194-204.

Lopez, Cecilio. 1965. Contributions to a comparative Philippine syntax. *Lingua* 15: 3-16.

Lyons, John. 1977. *Semantics.* Cambridge: Cambridge University Press.

McCloskey, James. 1986. Inflection and coordination in Modern Irish. *Natural Language and Linguistic Theory* 4: 245-281.

McCloskey, James, and Kenneth Hale. 1984. On the syntax of person-number inflection in Modern Irish. *Natural Language and Linguistic Theory* 1: 487-534.

Platero, Paul. 1982. Missing noun phrases and grammatical relations in Navajo. *International Journal of American Linguistics* 48: 286-305.

Platzack, Christer. 1987. The Scandinavian languages and the null-subject parameter. *Natural Language and Linguistic Theory* 5: 377-401.

Ramos, Teresita V. 1971. *Tagalog Structures.* Honolulu: The University Press.

Rizzi, Luigi. 1982. *Issues in Italian Syntax.* Dordrecht: Foris.

Sohn, Ho-min. 1975. *Woleaian Reference Grammar.* Honolulu: University of Hawaii Press.

Srinivasachari, K. 1988. *Learn Hindi in 30 Days* (21st ed.). Madras: Balaji.

Strang, Barbara M. 1970. *A History of English.* London: Methuen.

Suarez, Jorge A. 1983. *The Mesoamerican Indian Languages.* Cambridge: Cambridge University Press.

Topping, Donald M. 1973. *Chamorro Reference Grammar.* Honolulu: University Press of Hawaii.

Truscott, John. 2000. The unity of nominal linking devices. *Proceedings of the National Science Council, ROC. Part C: Humanities and Social Sciences* 10: 142-155.

Underhill, Robert. 1976. *Turkish Grammar.* Cambridge, MA: MIT Press.

Welmers, William E. 1973. *African Language Structures.* Berkeley, CA: University of California Press.

Yallop, Colin. 1982. *Australian Aboriginal Languages.* London: Andre Deutsch.

John Truscott
Department of Foreign Languages and Literature
National Tsing Hua University
Hsinchu, Taiwan

Truscott, John

truscott@mx.nthu.edu.tw