C.-T. James Huang

On the Distribution and Reference of Empty Pronouns

1. Posing the Problem

1.1. Introduction

In a recent paper John R. Ross (1982) suggests that languages may be classified using Marshall McLuhan's (1964) "hot-cool" division of the media. According to McLuhan, a medium is "hot" if the communication process involves little or no audience participation, and "cool" if active audience participation is required. A movie is thus a "hot" medium in that most of what it is intended to convey is presented before the viewer's eyes, whereas the telephone is a "cool" medium in that the success of communication depends upon considerable participation by the hearer. Commercial TV programs may be relatively "hot" since communication requires relatively little effort on the part of the audience, but an Oriental painting, or Russell's *An Inquiry into Meaning and Truth,* will be "cool" since full appreciation of their messages requires considerable effort on the part of the viewer or reader. Ross suggests that the same analogy may be extended to classifying languages on the basis of the explicitness with which they express certain anaphoric elements. For example, English may be said to be a "hot" language because pronouns cannot in general be omitted from grammatical sentences, and the information required to understand each sentence is largely obtainable from what is overtly seen and heard in it. On the other hand, Chinese may be said to be a very "cool" language in that such pronouns are usually omissible (and are often more naturally omitted) from grammatical sentences, and understanding a sentence requires some work on the reader's or the hearer's part, which may involve inference, context, and knowledge of the world, among other things. Other languages can be depicted as having a status somewhere between these two extremes, allowing more freedom than the "hot" languages, but less

An abbreviated earlier version of this article was presented at the second West Coast Conference on Formal Linguistics, February 1983. Slightly different versions have also been presented at colloquium lectures at MIT, USC, Cornell University, the University of Arizona, the University of Hawaii, Seoul National University, and National Taiwan Normal University. I am very grateful to the many people who offered me their helpful discussions and generous comments at various stages. Among the people who provided the most help are Adrian Akmajian (who I now regret cannot see the final version), Wynn Chao, Jim Gair, Jeanne Gibson, Irene Heim, Jim Higginbotham, Rich Janda, Osvaldo Jaeggli, Mamoru Saito, Jeffrey Tung, and Dong Whee Yang. The suggestions of two anonymous LI reviewers are also gratefully acknowledged.

Linguistic Inquiry, Volume 15, Number 4, Fall 1984
0024-3892/84/040531-44 $02.50/0

than the "cool" ones, for the use of empty or zero pronouns. For the sake of concreteness, let us look at a few examples in terms of this "hot-cool" descriptive parameter.

First of all, English and French are among the "hot" languages. In English, for example, an empty pronoun may be used only in the subject position of a nontensed clause, but generally nowhere else. This is shown by the following sets of sentences:

(1) a. John promised Bill that he would see Mary.
b. John promised Bill that Mary would see him.
c. John promised Bill that he would see him.
d. John preferred his seeing Mary.
e. John preferred Mary's seeing him.

(2) a. John promised Bill [*e* to see Mary].
b. John preferred [*e* seeing Mary].

(3) a. *John promised Bill that [*e* would see Mary].
b. *John promised Bill that [Mary would see *e*].
c. *John promised Bill [Mary to see *e*].
d. *John preferred [Mary's seeing *e*].
e. *John promised Bill that [*e* would see *e*].
f. *John promised Bill [*e* to see *e*].

An empty pronoun may occur only as the subject of an infinitival clause, as in (2a), or of a gerundive clause, as in (2b), but nowhere else. It cannot occur as the subject of a tensed clause (3a) or as an object at all (3b–f). This restriction seems to have nothing to do with semantic or pragmatic factors. This is particularly clear in the following discourse. Although the reference of an otherwise omitted pronoun is perfectly clear, omission is prohibited:

(4) Speaker A: Did John see Bill yesterday?
Speaker B: a. Yes, he saw him.
b. *Yes, *e* saw him.
c. *Yes, he saw *e*.
d. *Yes, *e* saw *e*.
e. *Yes, I guess *e* saw *e*.
f. *Yes, John said *e* saw *e*.

As is well known, many languages do not have such a strict restriction on the distribution of a zero pronoun. A number of languages are known to allow a zero pronoun in the subject position of a tensed clause (in addition to that of a tenseless clause), though not in the object or nonsubject position. Italian and Spanish are such "medium-hot" languages. In Spanish, for example, a subject pronoun may be omitted from a grammatical tensed clause (though it need not be); an object clitic may not. Thus, although both (5a) and (5b) are well-formed with a pronoun or clitic overtly present, the use of a zero pronoun results in the contrast shown in (6):

(5) a. José sabe [$_{S'}$ que él ha sido visto por María].
José know that he has been seen by María
'José knows that he has been seen by María.'

b. José sabe [$_{S'}$ que María lo ha visto].
José know that María him has seen
'José knows that María has seen him.'

(6) a. José sabe [$_{S'}$ que *e* ha sido visto por María].
José know that has been seen by María
'José knows that [he] has been seen by María.'

b. *José sabe [$_{S'}$ que María *e* ha visto].
José know that María has seen
'José knows that María has seen [him].'

In the third, or "cool," type of language, even sentences of the form represented by (6b) are perfectly grammatical. In fact, all of the sentences that we have starred up to now are grammatical in such languages, if they are uttered in appropriate contexts. In the Chinese discourse below, for example, all of speaker B's answers are acceptable:[1]

(7) Speaker A: Zhangsan kanjian Lisi le ma?
Zhangsan see Lisi LE Q
'Did Zhangsan see Lisi?'

Speaker B:
a. ta kanjian ta le.
he see he LE
'He saw him.'

b. *e* kanjian ta le.
'[He] saw him.'

c. ta kanjian *e* le.
'He saw [him].'

d. *e* kanjian *e* le.
'[He] saw [him].'

e. wo cai [*e* kanjian *e* le].
I guess see LE
'I guess [he] saw [him].'

f. Zhangsan shuo [*e* kanjian *e* le].
Zhangsan say see LE
'Zhangsan said that [he] saw [him].'

Note the sharp contrast in acceptability between the Chinese discourse (7) and the English discourse (4). Besides Chinese, Japanese and Korean are well-known languages

[1] The *Pinyin* system of transliteration is used throughout. Some symbols used in the glosses are: *LE*, the perfective or inchoative aspect marker; *BA*, the preposition marking a preverbal object; *DE*, the modifier marker that occurs at the end of any prenominal modifier. For the reader not familiar with the romanization, the following information may be useful in giving a sense of how the Chinese examples sound: *x* is a voiceless palatal fricative, *q* is an aspirated palatal affricate, and *zh* is an unaspirated retroflex affricate.

that exhibit this maximal freedom for the use of zero pronouns. Other such languages include Imbabura Quechua and Portuguese. The following data from Imbabura Quechua are due to Cole (1982):

(8) Speaker A: nuka mishki-ta miku-rka-ngui-chu?
my candy-acc eat-past-2-Q
'Did [you] eat my candy?'
Speaker B: ari, miku-rka-ni.
yes eat-past-1
'Yes, [I] ate [it].'

Examples (9) and (10) from Portuguese, provided by Raymond Moody (personal communication), contrast sharply with the earlier Spanish examples:

(9) a. José sabe [$_{S'}$ que ele viu Maria].
José know that he saw Maria
'José knows that he saw Maria.'
b. José sabe [$_{S'}$ que Maria o viu].
José know that Maria him saw
'José knows that Maria saw him.'

(10) a. José sabe [$_{S'}$ que *e* viu Maria].
José know that saw Maria
'José knows that [he] saw Maria.'
b. José sabe [$_{S'}$ que Maria *e* viu].
José know that Maria saw
'José knows that Maria saw [him].'

1.2. A First Statement of the Problem

We have seen that there are three types of languages: "hot," "medium," and "cool," according to the extent to which they allow the use of a zero pronoun. A natural question that arises in our investigation of Universal Grammar and linguistic typology is what parameter or parameters of UG it is that enable languages to differ precisely in the way they do. Recent works in generative grammar have devoted considerable attention to this problem, namely, formulating what has come to be known as the *Pro-Drop Parameter* or the *Null Subject Parameter* (see Perlmutter (1971), Borer (1983), Chomsky (1981), Chomsky and Lasnik (1977), Jaeggli (1982), Taraldsen (1978), among many others). One important type of explanation that has been proposed to distinguish between "hot" languages like English and French and "pro-drop" languages like Italian and Spanish is based upon the idea of recoverability and the observation, due to Taraldsen (1978), that the possibility of pro drop in a language often correlates with the existence in it of a rich inflectional morphology, in particular a rich system of agreement. According to this theory, as assumed in Chomsky (1981; 1982), Italian and Spanish allow a pronoun

to drop from the subject position of a tensed clause because there is a rich system of verb-subject agreement in these languages. The agreement marking on a verb is rich enough to determine, or recover, the content (i.e. reference) of a missing subject; therefore, such a missing subject is allowed. On the other hand, the agreement systems of English and French are somewhat degenerate, and the agreement marking on a verb is too meager to identify the content of a missing subject; therefore, such pronouns may not drop. Furthermore, since neither the Italian type nor the English type of language exhibits any verb-object agreement, no object pronoun may drop in any of these languages.

This mode of explanation appears to be quite plausible, and is further supported by an important piece of evidence from Pashto, a split ergative language spoken in Afghanistan.[2] In sentences expressing events of the present, Pashto uses a rich agreement system that is typical of nominative-accusative languages: the verb agrees with the subject in both transitive and intransitive sentences. In sentences expressing past events, however, the Pashto agreement system is ergative: the verb agrees with the subject if intransitive, but with the object if transitive. The accusative present agreement system is illustrated in (11a–b), and the ergative past agreement system is illustrated in (12a–b):

(11) a. Jān ra-z-i.
John DIR-come-3msg
'John comes.'

b. zə maṇa xwr-əm.
I apple eat-1msg
'I eat the apple.'

(12) a. Jān ra-ğ-ay.
John ASP-come-3msg
'John came.'

b. ma maṇa wə-xwar-a.
I apple PRF-eat-3fsg
'I ate the apple.'

In both the intransitive (11a) and the transitive (11b), the verb agrees with the subject. This also happens with the intransitive (12a). In the transitive (12b), however, the verb agrees with the third feminine singular object NP 'the apple'. Consider now what happens when pro drop occurs in this language. According to the theory based on the "Taraldsen generalization," only subjects may drop from sentences of the type represented by (11a–b) and (12a) and only objects may drop from those of the type (12b). This is in fact what happens: precisely the argument with which the verb agrees may drop, and only such

[2] I am indebted to Farooq Babrakzai for help with the Pashto data.

an argument. This is particularly clear from the contrast between (13b) and (14b):

(13) a. *e* ra-z-i.
DIR-come-3msg
'[He] comes.'

b. *e* maṇa xwr-əm.
apple eat-1msg
'[I] eat the apple.'

(14) a. *e* ra-ğ-ay.
DIR-come-3msg
'[He] came.'

b. ma *e* wə-xwar-a.
I PRF-eat-3fsg
'I ate [it (fem.)].'

Crucially, the object 'the apple' may not drop from (11b), and the subject 'I' may not drop from (12b):

(15) *zə *e* xwr-əm.
I eat-1msg
'I eat [?].'

(16) **e* maṇa wə-xwar-a.
apple PRF-eat-3fsg
'[?] ate the apple.'

The content of the object of (15) and of the subject of (16) must be expressed fully either as in (11b) and (12b) or with an object clitic and a subject clitic, respectively, as in (17) and (18):

(17) zə-ye xwr-əm.
I-clitic eat-1msg
'I eat it.'

(18) maṇa-me wə-xwar-a.
apple-clitic PRF-eat-3fsg
'I ate the apple.'

Pashto thus provides very interesting evidence for the theory of pro drop as proposed in Chomsky (1981; 1982), where the possibility of dropping a pronoun is closely tied to the presence of agreement.[3]

[3] Further evidence for the theory may be derived from languages in which the verb agrees with more than one argument. In Swahili, the verb agrees with the subject and the object, and both these NPs may drop. In Georgian (Stephen Anderson (personal communication)), the verb agrees with every argument, and everything may drop. In Arabic (Jelinek (1983)), a subject may drop only when a given finite verb shows agreement, but not when it does not. McCloskey and Hale (1983) show that in Irish, a verb, a preposition, or a head noun may or may not be inflected for person-number of the subject, the prepositional object, or the possessive (respectively). When it is so inflected (occurring in the "synthetic form"), the related argument may (and must) drop, and when it is not so inflected (occurring in the "analytic form"), pro drop is prohibited. The

This theory, however, runs into difficulty when we consider the third type of language, such as Chinese, Japanese, and Korean, which have no system of verb-subject or verb-object agreement. A theory based solely on the "Taraldsen generalization" would predict that such languages allow neither zero subject nor zero object pronouns. However, as we have seen, the situation is precisely the opposite, for these languages apparently allow pro drop even more freely than those with rich agreement systems.

We are thus faced with the paradox of having a theory that must be right for some reasons but must be wrong for others. The problem is how to resolve this paradox. One plausible solution could be to assume that the principle based upon recoverability and agreement may itself be parametrized across languages. Certain languages are required to rely on agreement and meet the condition of recoverability; others simply are not required to do so. For reasons that should become clear shortly, however, I feel that this simple way of tackling the problem will not lead to a satisfactory and illuminating solution. To see the problem in the right way, we need to consider the facts of the "cool" languages in more detail, as we will do in the next section.

1.3. Restrictions on the Zero Object

Consider the following sentences in Chinese:[4]

(19) a. *e* lai-le.
come-LE
'[He] came.'

b. Lisi hen xihuan *e*.
Lisi very like
'Lisi likes [him] very much.'

c. Zhangsan shuo [*e* bu renshi Lisi].
Zhangsan say not know Lisi
'Zhangsan said that [he] did not know Lisi.'

d. Zhangsan shuo [Lisi bu renshi *e*].
Zhangsan say Lisi not know
'Zhangsan said that Lisi did not know [him].'

Compare these with the following sentences in English, where an overt pronoun appears

person-number marking may be taken as Agr marking, and Irish may be seen as providing further support for the Agr-based theory. Because pro drop is obligatory in environments of the "synthetic forms," however, McCloskey and Hale also allow for the possibility that the markings are clitics, or incorporated pronouns, rather than Agr. Note that the fact is clear in Pashto. The marking on a given verb may cooccur with a lexical NP, as shown in (11) and (12), and is therefore agreement rather than an incorporated pronoun. This view about Pashto is further supported by the fact that the language has a separate set of clitics, as shown in examples (17) and (18).

[4] Although *e* in (19a) is translated as 'he', this is not the only possible interpretation. Depending on the context, it may be 'I', 'she', etc. The same remark applies to certain other examples throughout.

instead of an empty category (henceforth, EC):

(20) a. He came.
b. Bill saw him.
c. John said that he knew Bill.
d. John said that Bill knew him.

The EC in (19a) refers to someone whose reference is understood in discourse, on a par with *he* in (20a). Likewise, the reference of the EC in (19b) is fixed outside of the sentence, on a par with *him* in (20b). In (19c) the embedded subject EC may refer either to the matrix subject *Zhangsan* or to someone whose reference is fixed outside of the sentence. This is, again, exactly as in (20c), where the embedded subject pronoun may refer either to *John* or to someone whose reference is distinct from *John*.

We have seen a striking similarity in distribution and reference between an overt pronoun in English and a zero pronoun in Chinese. When we turn to (19d) and (20d), an important difference emerges between them. In (20d) the pronoun *him* is free in reference: it may refer to the matrix subject *John*, or it may refer to someone whose reference is distinct from *John*. In (19d), however, the embedded object EC may refer only to someone whose reference is fixed outside of the entire sentence, but not to the matrix subject *Zhangsan*.[5] To convey the ambiguous message that Lisi knew either Zhangsan or someone else, the EC would have to be replaced by an overt pronoun:

(21) Zhangsan shuo Lisi bu renshi ta.
Zhangsan say Lisi not know him
'Zhangsan said that Lisi didn't know him.'

There is, then, some restriction on the reference of an EC if it occurs as an object, but not if it occurs as a subject. In other words, for certain interpretations, the distribution of an object EC is more limited than that of a subject EC. More examples showing this subject-object asymmetry are given below. In each case, consider only the interpretation according to which the embedded EC is bound by the matrix subject:

(22) a. Zhangsan$_i$ xiwang [e_i keyi kanjian Lisi].
Zhangsan hope can see Lisi
'Zhangsan$_i$ hopes that [he$_i$] can see Lisi.'
b. *Zhangsan$_i$ xiwang [Lisi keyi kanjian e_i].
Zhangsan hope Lisi can see
'Zhangsan$_i$ hopes that Lisi can see [him$_i$].'

[5] The relevant point is that the empty category in (19d), when construed as the *object* of the embedded verb, cannot referentially depend upon the matrix subject *Zhangsan*. The EC may be interpreted as referring to *Zhangsan* if it is construed as the subject of the embedded verb 'know'. The *Lisi* immediately preceding the embedded verb will, in this case, have to be construed as a preposed object. This situation is the same as indicated below, where the NP 'this book' must be construed as a preposed object:

(i) Zhangsan shuo [zheben shu *e* hen xihuan].
Zhangsan say this book very like
'Zhangsan said that this book, [he] likes very much.'

Note that in order to see the contrast between (19c) and (19d) clearly, some care is needed to ensure that they are considered in pragmatically neutral contexts.

(23) a. Zhangsan$_i$ zhidao [e_i mei banfa shuifu Lisi].
Zhangsan know no method persuade Lisi
'Zhangsan$_i$ knows that [he$_i$] cannot persuade Lisi.'
b. *Zhangsan$_i$ zhidao [Lisi mei banfa shuifu e_i].
Zhangsan know Lisi no method persuade
'Zhangsan$_i$ knows that Lisi cannot persuade [him$_i$].'

The point being made here deserves some emphasis and clarification. To be more precise, we should say that the relevant contrast between (19c) and (19d), and between the (a) and (b) sentences of (22) and (23), concerns what a given EC can take as its *antecedent,* not simply what it can corefer with. It has sometimes been observed that given an appropriate context it is possible to construe the object EC in (19d), (22b), or (23b) as referring to the matrix subject. In particular, if (19d) is uttered as a reply to the question *Who didn't know Zhangsan?*, it is entirely natural to construe the EC as referring to *Zhangsan*. Thus, consider the following discourse:

(24) Speaker A: Shei kanjian-le Zhangsan?
who see-LE Zhangsan
'Who saw Zhangsan?'
Speaker B: Zhangsan shuo Lisi kanjian-le *e*.
Zhangsan say Lisi see-LE
'Zhangsan said Lisi saw him.'

The EC in speaker B's reply can certainly refer to *Zhangsan*. The relevant claim that I am making, however, is just that in a sentence like (19d) the object EC cannot *referentially depend* upon the matrix subject as its antecedent. Thus, in (24), although the EC in speaker B's reply may corefer with the matrix subject, the claim is that the latter is not its antecedent. Rather, the antecedent is the occurrence of *Zhangsan* in speaker A's question. (For more discussion of the notion of referential dependency, see Evans (1980).)

It is therefore important that in order to see the contrast between (19c) and (19d) and similar pairs of sentences, one should consider them in contexts in which pragmatic or discoursal factors are reduced to the minimum. The relevant point concerning (19c) and (19d) is that the contrast obtains when both are uttered without a context. If, for example, someone suddenly opens the door and says (19c), there is no difficulty in assuming that the EC refers to *Zhangsan*, but if the visitor says (19d), the usual assumption is that the EC refers to someone else, and the most likely response is "Zhangsan said Lisi didn't know who?" Note that if each of the ECs in (19c) and (19d) is replaced by a lexical pronoun, then the contrast disappears entirely.[6]

[6] What (24) illustrates is that a grammatical point may be obscured by discoursal factors. The following example illustrates the same point:

(i) xiaotou yiwei meiyou ren kanjian *e*, na le dongxi jiu pao.
thief think no man see take LE thing then run
'The thief thought no one saw [him], so he took the things and ran.'

What we have observed about the distribution of certain ECs in Chinese also holds true to some extent in Japanese. Contrasts of the following sort have been observed by Kuroda (1965):

(25) a. dare-ga$_i$ [e_i Bill-o nagutta] to itta ka?
who Bill hit that said Q
'Who$_i$ said that [he$_i$] hit Bill?'

b. *dare-ga$_i$ [Bill-ga e_i nagutta] to itta ka?
who Bill hit that said Q
'Who$_i$ said that Bill hit [him$_i$]?'

(26) a. John-wa$_i$ [e_i siken-ni otita] no o mada siranai.
John exam failed that yet not-know
'John$_i$ still doesn't know that [he$_i$] failed the exam.'

b. *John-wa$_i$ [Bill-ga e_i settokusuru] to omotte iru.
John Bill persuade that think
'John$_i$ thinks that Bill will persuade [him$_i$].'

The asymmetry shown in (25)–(26) is the same asymmetry that we have observed in Chinese.[7] It can also be observed in Korean. In the following sentence (from D. W. Yang (personal communication)), the embedded clause contains a subject EC and an object EC:

(27) John-i [*e e* po-əss-ta-ko] malha-əss-ta.
John-SUBJ see-PAST-DECL-COMP say-PAST-DECL

In this sentence the object EC can refer to the matrix subject 'thief', since this is pragmatically the most natural way to interpret the sentence. The reference of the EC in this sentence is pragmatically inferred, but not grammatically determined. The situation with (19d), (22b), and (23b) is different, since pragmatic considerations do not force us to interpret the embedded object EC in each of them as being disjoint from its matrix subject. This is evidenced by the fact that substituting an overt pronoun for each of these ECs immediately results in an interpretation according to which no disjointness of reference is required. A relevant point to consider here is that in situations such as those represented by (i) and (24), pragmatics appears to "override" grammar in Chinese, so that (i) and speaker B's reply in (24) are not ruled out as unacceptable utterances, though comparative sentences in English, say, are still unacceptable: **The thief thought that no one would see e, so he . . .* .In purely descriptive terms the difference between Chinese and English is in the extent to which pragmatics may override grammar (cf. Comrie (1983)). An LI reviewer has pointed out correctly that this difference cannot surely be taken as some sort of primitive, since it must be owing to the design features of the language itself that pragmatics can override grammar. I will assume that this difference will have to be derived from the more general parameter distinguishing between "discourse-oriented" and "sentence-oriented" languages, as will be indicated in section 2.

What I wish to insist upon in this discussion is that facts that occur in special pragmatic contexts should not be taken to deny the existence of certain grammatical facts that are often observable only in pragmatically neutral contexts.

[7] In (26) the NP *John* is marked as the topic by *-wa*. As will become clear later, what is relevant is that this topic NP is also the matrix subject. Like the other examples we have seen, the sentence is rejected because the embedded EC is interpreted as bound by the matrix subject. If we assume that there is a matrix subject EC immediately following, and coindexed with, the topic, then (26) is rejected because the embedded EC is bound by this empty matrix subject. That there is no special restriction against having the embedded EC bound by the topic per se is shown by the fact that, if the topic is distinct from the matrix subject, it can bind the embedded EC. Compare (26) with the following:

(i) John-wa$_i$, minna-ga [Bill-ga e_i settokusuru] to omotte iru.
John everyone Bill persuade that think
'John$_i$, everyone thinks that Bill will persuade e_i.'

The matrix subject *John* may be understood to be saying that he saw someone else, but not that someone else saw him. This means that of the two embedded ECs, only the subject but not the object may be bound by the matrix subject *John*.

A similar observation may be made about the two Portuguese sentences (10a) and (10b). The following sentences in Brazilian Portuguese (Chao (1983b; personal communication)) show the same asymmetry. In each (a) sentence, the embedded subject EC may be bound by the matrix subject (or by a discourse topic), but in each (b) sentence, the embedded object EC may not be bound by the matrix subject:[8]

(28) a. João$_i$ disse que e_i viu o Pedro.
João said that saw Pedro
'João said that he saw Pedro.'

b. *João$_i$ disse que Pedro viu e_i.
João said that Pedro saw

(29) a. João$_i$ sabe que e_i gostaria de conhecer a Maria melhor.
João know that would-like know Maria better
'João knows that he would like to know Maria better.'

b. *João$_i$ sabe que a Maria gostaria de conhecer e_i melhor.
João know that Maria would-like know better

In fact, an embedded object EC cannot be bound to a matrix object either:

(30) a. João disse à Maria$_i$ que gostaria de conhece-la$_i$ melhor.
João said Maria that would-like know-her better
'João told Maria that he would like to know her better.'

b. *João disse à Maria$_i$ que gostaria de conhecer e_i melhor.
João said Maria that would-like know better

It should be obvious by now, from the asymmetry we have observed in several languages, that the problem posed at the end of section 1.2 cannot be entirely solved by stipulating that some languages must obey certain pro-drop conditions while others need not. Such a simple solution does not lead to an explanation of the asymmetry. Let us examine this asymmetry more closely now, with a view to properly sorting out the problem and finding a solution.

1.4. Sorting Out the Problem

We have seen that in a language like Chinese, an object EC may not be bound by a matrix argument, though it may be bound by some NP whose reference is fixed in discourse. Its reference must, in other words, be the discourse topic, someone or something that a given discourse is about. A closely related fact that should be noted now is that, when a topic NP appears in addition to a matrix subject, an embedded object EC is most naturally interpreted as bound by that topic. The following sentences are identical in surface form to (19d), (22b), and (23b), except that they each contain a topic, *neige*

[8] For more discussion of null NPs in Brazilian Portuguese, see Chao (1981; 1983a,b).

ren 'that man', whose reference is distinct from the matrix subject *Zhangsan*. As shown, each sentence is well-formed with the object EC most naturally interpreted as bound by the topic:

(31) neige ren$_i$ Zhangsan shuo [Lisi bu renshi e_i].
that man Zhangsan say Lisi not know
'That man$_i$, Zhangsan said Lisi didn't know e_i.'

(32) neige ren$_i$, Zhangsan xiwang [Lisi keyi kanjian e_i].
that man Zhangsan hope Lisi can see
'That man$_i$, Zhangsan hopes that Lisi will be able to see e_i.'

(33) neige ren$_i$, Zhangsan zhidao [Lisi mei banfa shuifu e_i].
that man Zhangsan know Lisi no method persuade
'That man$_i$, Zhangsan knows that Lisi won't be able to persuade e_i.'

There is a clear similarity between the set of sentences (19d), (22b), (23b) and the set (31)–(33). In both cases the object EC refers to a topic, the only difference being that in (19d) and the (b) sentences of (22)–(23) the topic is not overtly present in a sentence (but can be inferred in discourse), whereas in (31)–(33) the topic is overtly expressed. Within our conceptual framework, what is not overtly present in a given structure is what we have been representing as an EC. This naturally suggests that (19d) and the (b) sentences of (22)–(23) may be assimilated to (31)–(33) in structure, each with an empty topic. (19d) may be represented as (34), with an empty topic binding the embedded object EC, on a par with (31) with a lexical topic:

(34) (= (19d)) [$_{\text{Top}}$ e_i], [Zhangsan shuo [Lisi bu renshi e_i]].
Zhangsan say Lisi not know
'*[Him$_i$], Zhangsan said that Lisi didn't know e_i.'

Note that, seen in this light, what we have been assuming to be an object EC is not exactly an empty pronoun in the usual sense of the term. Within earlier models of generative grammar, there is simply no object gap in (31). The embedded object has been topicalized and appears in sentence-initial position, and there is nothing "missing" in the sentence. Similarly, within traditional analyses, one could say, given the facts presented here, that there is really no object gap in (34) (= (19d)). What is really missing is the topic—that is, an object is topicalized first before it is deleted from topic position. On the other hand, a subject EC of the type indicated below is a genuine zero pronoun, since these do not involve any movement at all:

(35) John$_i$ tried [e_i to come].

(36) Zhangsan$_i$ shuo [e_i mingtian yao lai].
Zhangsan say tomorrow want come
'Zhangsan$_i$ said that [he$_i$] wants to come tomorrow.'

Within the Government and Binding (GB) framework of Chomsky (1981; 1982), the difference between movement and base-generation of an EC is trivialized, and the identity of an EC is functionally determined according to the following principle (Chomsky

(1981, 330)):[9]

(37) a. An EC is a pronominal if and only if it is free or locally bound by an element with an independent thematic role, and a nonpronominal otherwise.
b. A nonpronominal EC is an anaphor if and only if it is locally A-bound, and a variable if locally Ā-bound.

According to (37), then, the embedded subject EC in (19c) may be a pronominal, since it may be locally bound by the matrix subject *Zhangsan*, which has its own thematic role as agent of the matrix verb 'said'. The object EC in (19d) (= (34)) and (31)–(33), on the other hand, cannot be a pronominal, since it is locally bound by a topic, which is in a non-A, nonthematic position. Such an EC is by (37) a variable.

We have argued that an object EC cannot be a pronominal, though it can be a variable, by showing that it cannot be locally A-bound by a matrix argument, though it can be locally Ā-bound by a topic. Another piece of support for this claim comes from examples where such an EC can only be locally Ā-bound by the head of a relative clause but not by an NP in argument position. Consider (38):

(38) Li Xiaojie hai zhao-bu-dao [yige [*e* xinzhong xihuan *e* de] nanren].
Li Miss still can't-find one in-heart like DE man
'Miss Li still cannot find a man who she loves in her heart.'
Not: 'Miss Li still cannot find a man who loves her in his heart.'

There are two ECs in the relative clause in (38). If each EC is bound by an NP within the main clause, then the sentence has two potential interpretations: either (a) the subject EC is bound by the subject 'Miss Li' (an A-binder) and the object bound by the head of the relative clause 'a man' (an Ā-binder), or (b) conversely. According to the (a) reading, the subject EC is a pronominal and the object EC a variable; and according to the (b) reading, the subject is a variable and the object a pronominal. As indicated in the translation, this sentence is not ambiguous; only the first reading is available. This lack of ambiguity is a direct consequence of the generalization that an object EC cannot be a pronominal. The contrast below shows exactly the same point. The contrast occurs because the verb *jia* 'marry (off)' in (39) requires a female subject, and *qu* 'marry (in)' requires a male subject:[10]

(39) Li Xiaojie hai zhao-bu-dao [yige [*e* keyi jia *e* de] nanren].
Li Miss still can't-find one can marry DE man
'Miss Li still cannot find a man who [she] can marry.'

[9] α binds β iff α and β are coindexed and α c-commands β. A category is free iff it is not bound. α A-binds β iff α binds β and α is in an argument position (subject, object, etc.), and α Ā-binds β iff α binds β and α is in an operator, nonargument position (Comp, Top, etc.). Chomsky distinguishes two kinds of pronominal ECs: the pronominal anaphor PRO (ungoverned) and the pronominal nonanaphor pro (governed).

[10] Against the unacceptable sentence (40), an LI reviewer suggests the following as a possible counterexample:

(i) neige gu-er$_i$ zhao-bu-dao [[*e* yuanyi fuyang e_i de] ren].
that orphan cannot-find willing adopt DE man
'That orphan$_i$ cannot find someone who is willing to adopt [him$_i$].'

(40) *Li Xiaojie hai zhao-bu-dao [yige [*e* keyi qu *e* de] nanren].
Li Miss still can't-find one can marry DE man
'Miss Li still cannot find a man who can marry [her].'

What we see here can again be observed in Japanese. The unambiguous sentence (41), from Ohso (1976, 6), is on a par with (38):

(41) John-wa [*e e* aisite mo inai] onna to kekkonsita.
John love even doesn't woman married
'John married a woman who he doesn't even love.'
Not: 'John married a woman who doesn't even love him.'

The same point can be made with free relatives containing more than one gap. Consider first of all the following two sentences:

(42) [[*e* mai fangzi de]] gen [[*e* zu fangzi de]] dou hao.
buy house DE and rent house DE all good
'Both the one who buys houses and the one who rents houses are good.'

(43) [[ni mai *e* de]] gen [[wo zu *e* de]] dou hao.
you buy DE and I rent DE all good
'Both what you bought and what I rented are good.'

Suppose that a free relative is a relativized construction with an empty head (cf. Teng (1979), Kitagawa and Ross (1982)); then each EC in (42) and (43) is nothing more than a relativized argument, as a variable bound by the (empty) head. Thus, the first free relative in (42) has the following structure, where the empty head is marked by Ø:[11]

(44) [$_{NP}$[$_{S'}$ *e* mai fangzi de] Ø].
buy house DE

Each of the free relatives has exactly one EC in argument position: in (42) each EC is in subject position, and in (43) each EC is in object position. Since there is exactly one EC in each relative clause, this EC is interpreted as relativized, as indicated in the translations above. Now consider a sentence in which a free relative contains both a

---

The embedded object EC can apparently refer to the matrix subject. Because of the use of the embedded verb 'adopt', apparently the reference of the EC can be pragmatically inferred, and this sentence need not be taken as a counterexample, since it does not represent a pragmatically neutral context. The question does arise, however, why (40) is not also acceptable with the object EC also pragmatically inferred as referring to 'Miss Li'. My own suggestion is that (40) is unacceptable because of the existence of the verb *jia* 'marry off' in addition to *qu* 'marry in'. That is, the unavailable reading of (40) can be expressed by the readily available (39), and this eliminates any motivation for resort to a pragmatic principle in violation of a grammatical principle.

[11] Chinese has two ways to render free relatives like *what you saw*: one by relativization, as shown by the examples in the text, with an empty head, and the other by the use of *wh*-quantifiers, as in constructions like (i):

(i) ni yao shenme, wo dou gei ni.
you want what I all give you
'Whatever you want, I will give to you.'

subject and an object gap (cf. Li and Thompson (1981, 578)):

(45) [[*e* mai *e* de]] gen [[*e* zu *e*]] dou hao.
buy DE and rent all good
'What one buys and what one rents are both good.'

As indicated by the translation, each subject gap in (45) is interpreted as 'one', as a pronoun having arbitrary reference, and each object gap is taken to be the relativized argument. Given that both the subject and the object may be relativized (as indicated by (42) and (43)), a sentence like (45) might be expected to have a second reading, according to which the object gap would be an arbitrary pronoun and the subject gap relativized. According to this, the sentence would mean 'The one who buys things and the one who rents things are both good'. Indeed, one might even expect the sentence to be four-ways ambiguous, given that each free relative contained in it may be two-ways ambiguous. In fact, however, the sentence is not ambiguous at all. It has only the interpretation indicated. This is precisely the one situation, out of the four possibilities, where neither object EC in (45) is taken as a pronominal.[12] Again, the same point may be observed in Japanese:

(46) [*e e* katta] no wa omoshiroi desu; [*e e* karita] no wa omoshirokunai.
bought one Top interesting is rented one Top uninteresting
'What was bought is interesting; what was rented is uninteresting.'

Summarizing, I have claimed here that "cool" languages like Chinese do not differ from the "noncool" languages in allowing an object zero pronoun. Rather, they also prohibit empty object pronouns. The difference between "cool" and "noncool" languages is now recast as the difference between allowing a *zero topic* binding a variable (as in (34)) and not allowing such a topic. The facts we have seen so far may be summarized as shown in table 1.[13]

If this is the correct arrangement of the facts reviewed so far, then it seems that three distinct questions may be asked: First, what determines the distribution of a zero

[12] The asymmetry described in the text obtains crucially, again, only in a pragmatically neutral context, when one of the ECs is taken to be generic or arbitrary in reference (and the other taken to be relativized). In the following sentence the object EC need not be taken as relativized:

(i) neiben shu, [*e* kan-guo *e* de] bu shao.
that book read-has DE not few
'That book, the ones that have read [it] are not few.'

As indicated, the subject EC may be taken to be relativized, and the object EC taken to refer to the topic 'that book'. The same may be true even if the topic is not overtly indicated as 'that book' in the sentence, but simply understood to be the discourse topic—that is, when there is a "zero topic" as in (34). Note that in these cases, the object EC is still not a pronominal by definition. It is bound to the topic, and, like a relativized EC, it is a variable. The relevant point is that although a subject EC can be a pronominal having arbitrary reference (like the arbitrary PRO), an object EC cannot. (See Suñer (1983) for indications that there may be arbitrary pro in addition to arbitrary PRO.)

[13] Henceforth, I will concentrate on languages with no agreement or only subject-verb agreement, and will not consider languages like Pashto. It should be remembered that when I claim that an object EC cannot be a pronominal, I do not refer to languages showing verb-object agreement.

Table 1.

Types of ECs	"Hot" Languages	"Medium" Languages	"Cool" Languages
Zero subject (PRO) in tenseless clauses?	Yes	Yes	Yes
Zero subject (pro) in tensed clauses?	No	Yes	Yes
Zero object (pro)?	No	No	No
Zero topic?	No	No	Yes

topic, which distinguishes "cool" from "noncool" languages? Second, why is it that all the languages under investigation allow only zero subject pronouns, but exclude zero object pronouns? Third, what is the relevant factor that determines the degree to which a language allows a zero subject pronoun?

My eventual purpose in this article is to suggest a line of research that may lead to solving these problems. Before we move toward a possible solution, however, it is worthwhile to point out that the way I have arranged the facts so far not only seems maximally plausible, but also is supported by an independent piece of evidence from German, which I now present.

1.5. Pronoun Zap in German

As pointed out by Ross (1982), in spoken German one can drop either a subject or an object pronoun from a grammatical sentence. This produces examples of the following kind:

(47) a. Ich hab' ihn schon gesehen.
I have him already seen
'I saw him already.'

b. Hab' ihn schon gesehen.
have him already seen
'I saw him already.'

c. Hab' ich schon gesehen.
have I already seen
'I saw him/it/her/them already.'

(48) a. Ich trage die schon.
I wear them already
'I am wearing them already.'

b. Trage die schon.
I them already
'I am wearing them already.'
c. Trage ich schon.
wear I already
'I am wearing them/it already.'

Each of the (a) sentences retains both its subject and its object. In the (b) sentences the subject pronoun has been deleted, and in the (c) sentences the object pronoun has been deleted. An important restriction observed by Ross on the process of "Pronoun Zap" is that deletion of a given NP may take place only when the NP appears in sentence-initial—namely, topic—position. In the (b) and (c) sentences of (47) and (48), this is shown by the fact that the verb appears initially on the surface in each of these sentences. Given the well-known requirement of German that the verb must appear in second position, it is clear that the missing NP would be in first or topic position if it were present. In other words, (47b) and (47c), for example, have the following representations:

(49) *e* hab' ihn schon gesehen.
have him already seen
(50) *e* hab' ich schon gesehen.
have I already seen

Crucially, deletion cannot take place from a nontopic position. Compare the acceptable (47b–c) and (48b–c) with the unacceptable forms in (51)–(52). In each of these sentences, the first position is already lexically filled, which forces the EC to occur in the third position, where it is not allowed.[14]

(51) a. *Ihn hab' *e* schon gesehen.
b. *Ich hab' *e* schon gesehen.
(52) a. *Die trage *e* schon.
b. *Ich trage *e* schon.

A related restriction is that only one argument may drop per sentence. This may not be immediately obvious with sentences like (47), given that (53) is acceptable with

[14] Rich Janda has pointed out to me that the phenomenon described is not limited to noun phrase deletion. Phrases of other types may also be omitted, but again only from the topic position.

(i) Speaker A: Du mußt dein Bett machen.
you must your bed make
'You have to make your bed.'
Speaker B: *e* habe ich schon!
have I already
'I have already!'

It has been noted that this type of pronoun deletion is permitted only in informal speech. The relevant point distinguishing German from English is that even in comparable informal speech in English the same type of omission is not allowed.

no overt subject or object:

(53) Hab' schon gesehen.

Irene Heim (personal communication) has indicated to me, however, that (53) is acceptable only when its meaning corresponds to 'I understand' or 'I am aware' in English. (53) cannot be interpreted as 'I saw him'. This suggests that (53) may be more appropriately treated as intransitive, so that there is only one EC in it.[15] With a verb like *kennen* 'know (a person)', which cannot have an intransitive use, this one-gap-per-sentence restriction is clear. In the paradigm below, note crucially the ill-formedness of (54e):

(54) a. Ich hab' ihn schon gekannt.
I have him already known
'I already knew him.'
b. Ihn hab' ich schon gekannt.
c. *e* hab' ich schon gekannt.
d. *e* hab' ihn schon gekannt.
e. **e* hab' *e* schon gekannt.
f. *Ich hab' *e* schon gekannt.
g. *Ihn hab' *e* schon gekannt.

Note that the one-gap-per-sentence restriction may be seen as a simple consequence of the earlier restriction that an NP may be deleted only in topic position, given the usual assumption that there is only one topic position in German, the verb being always in second position.

These facts from German provide very interesting support for the proposed view of the facts arranged in table 1. German is like Chinese in that it allows an object NP to be missing in a grammatical sentence. However, what appears to be a zero object pronoun in German turns out to be a variable bound by a zero topic—it must first be topicalized before it is deleted from the topic position. This is exactly how I proposed to look at apparent zero object pronouns in Chinese. The two languages share the property of allowing a variable bound by a zero topic, although the evidence for this property is directly "visible" only in German, since only German has the "verb-second" requirement.

Another piece of support is derivable from the one-gap-per-sentence restriction in German, which distinguishes it from Chinese-type languages. The German requirement suggests that *every* NP to be deleted must be topicalized first: there is no genuine zero pronoun at all in a tensed clause in this language, only zero topics. On the other hand, I have indicated that in Chinese-type languages, an embedded *subject* EC may be a genuine zero pronoun (as indicated by the fact that *e* in (19c) may be A-bound by *Zhang-*

[15] Therefore, (51b) is, strictly speaking, not ungrammatical. It is well-formed, but has the meaning corresponding to 'I was already aware'.

san). I will propose shortly that there are two distinguishable parameters accounting for the facts under consideration. One distinguishes zero-topic languages like Chinese from non-zero-topic languages like English and Italian. The other distinguishes languages that allow zero subjects in tensed sentences from those that do not (the Pro-Drop Parameter). Logically, then, there can be four different types of languages. English and French are neither zero-topic nor pro-drop languages. Italian and Spanish are pro-drop, but not zero-topic, languages. Chinese, Japanese, Portuguese, etc., are both pro-drop and zero-topic languages. German appears to be an example of the fourth type: a zero-topic but non-pro-drop language. It thus provides important evidence for our theory by filling an otherwise existing peculiar gap in the proposed typological scheme.

2. Toward a Solution

I have devoted considerable space to the task of determining what the problems are. This rather lengthy discussion is justified by the often expressed truth that posing a problem right automatically solves half of it. Let us turn now to the other half of the task.

2.1. Occurrence of the Zero Topic

Consider the first problem—namely, what determines the difference between "cool" and "noncool" languages with respect to the distribution of zero topics. I would like to suggest that this difference may be derived from a more general typological parameter that has been proposed in the literature. Specifically, I refer to the work of Tsao (1977), who argues that languages like Chinese may be distinguished from languages like English by a parameter that he calls "*discourse-oriented* vs. *sentence-oriented.*" Chinese is a "discourse-oriented" language and English is a "sentence-oriented" language. Although Tsao is chiefly concerned with Chinese, his remarks can very well be extended to Japanese, Korean, and so forth. The motivation for establishing this typological parameter lies in the fact that languages that may be grouped as being "discourse-oriented" exhibit a clustering of distinctive properties that the "sentence-oriented" languages do not. Among the properties that Tsao enumerates in support of this parameter is the fact that discourse-oriented languages have a rule of *Topic NP Deletion,* which operates across discourse to delete the topic of a sentence under identity with a topic in a preceding sentence. The result of such a deleting process is formally a *topic chain.* Example (55) shows a topic chain that has undergone Tsao's Topic NP Deletion:

(55) [Zhongguo, difang hen da.] [*e*, renkou hen duo.] [*e*, tudi hen
China place very big population very many land very
feiwo.] [*e*, qihou ye hen hao.] [*e*, women dou hen xihuan.]
fertile climate too very good we all very like
'(As for) China, (its) land area is very large. (Its) population is very big. (Its) land is very fertile. (Its) climate is also very good. We all like (it).'

Each of the ECs above marks the site of a deleted topic. This is, of course, what we have been calling a zero topic. The idea behind Topic NP Deletion can be easily incorporated into an interpretive framework. We may assume that there is a rule of coindexation, in the discourse grammar of a discourse-oriented language (in the LF′ module of grammar following LF), which coindexes an empty topic node with an appropriate preceding topic. On the other hand, a sentence-oriented language has a less substantive discourse grammar in that it lacks this topic-chain interpretation rule, among others.

Another property of discourse-oriented languages is what Li and Thompson (1976a) call "*topic-prominence*." Discourse-oriented languages are more "topic-prominent," and sentence-oriented languages are more "subject-prominent," since topic is more of a discourse notion than subject, which is a syntactic notion. In a subject-prominent language like English, all sentences must have subjects—a property that Chomsky (1982, 10) describes as part of the Extended Projection Principle. This accounts for the presence of pleonastic elements like *it* and *there* in such languages. On the other hand, in a topic-prominent language like Chinese, Japanese, or Korean, structural subjects are not a basic requirement of the sentence, and these languages do not have such pleonastic elements. In such languages, furthermore, sentences of the form topic-comment abound and must count as basic forms in that they cannot be plausibly derived from other, "more basic" forms. An example of this well-known fact is the following sentence (Li and Thompson (1981, 96)):

(56) neichang huo, xingkui xiaofangdui lai de zao.
that fire fortunately fire-brigade come COMP early
'That fire, fortunately the fire brigade came early.'

One might plausibly assume that the basic nature of a topic-comment sentence in such a language is what gives rise to the possibility of allowing independent sentences each of which contains a "zero topic."[16]

A third fact supporting the parameter under consideration is that in a discourse-oriented, but not sentence-oriented, language, an anaphor may be discoursally bound. This observation, due to Yang (1983), is illustrated by the following discourse in Korean:

(57) Speaker A: John-i salam-ɨl ponae-əss-ni?
John-NOM man-ACC send-PAST-Q
'Did John send the man?'

[16] Gundel (1980) has independently proposed a view that is somewhat similar to that taken here. In particular, she suggests in very general terms that the more "topic-prominent" a language is, the more likely it is to drop a pronoun. She claims that this accounts for all cases of zero anaphora. This approach is defective in several respects. First, it ignores the relevance of agreement as demonstrated by the Pashto data. Second, it has nothing to say about the subject-object asymmetry observed here. Third, it says nothing about ECs in pro-drop (or "topic-prominent") languages that do not necessarily refer to topics. Fourth, it wrongly assumes that a zero pronoun must refer to a topic. Even with object ECs, what is relevant is not that such an EC must be bound by a topic, but that it must be a variable. An EC may be a variable by being bound by a topic, or by being relativized or otherwise Ā-bound.

Speaker B: ani, caki-ka cikcəp o-əss-ta.
no self-NOM in-person come-PAST-DECL
'No, self came in person.'

The answer given by speaker B is acceptable, and the nominative anaphor 'self' can easily be inferred as discoursally bound to *John* in the preceding sentence. Similar well-formed discourses can be constructed in Chinese.[17] Compare the following English discourse, in which the answer given by speaker B is unacceptable:

(58) Speaker A: Did John send the man?
Speaker B: *No, himself came.

The clustering of properties that we have seen to distinguish "discourse-oriented" from "sentence-oriented" languages shows that the parameter assumed here is independently motivated. Therefore, the difference between "cool" and "noncool" languages with respect to the distribution of a zero topic can be accounted for without additional cost.[18]

2.2. *Occurrence of a Genuine Zero Pronoun*

Let us turn now to the second and third problems and consider cases of an empty pronoun that occurs in an argument position rather than as an empty topic. The second problem concerns the fact that none of the languages we have studied allows a genuine zero object pronoun, though all languages allow a genuine zero subject pronoun to some extent. The problem is why this should be the case. In terms of earlier treatments in generative

[17] A sentence like *ziji lai le* 'self came' may be used as an answer to the question "Did John send someone?" Since *ziji* may also be used adverbially (meaning 'personally'), however, one might suspect that the sentence *ziji lai le* is really one with a zero subject and the adverbial *ziji*. (The case in Korean is clear, given the nominative case marking on *caki*.) It is possible to show that *ziji* can be a subject, though, as in *ziji zuotian lai le* 'self came yesterday'. The time adverb 'yesterday' normally occurs before a manner adverb, so 'self' should be a subject that normally occurs before the time adverb.

[18] There is still a nontrivial problem of execution that cannot be dealt with here. For example, the topic of a topic chain is not always introduced in the topic position in the first sentence of the chain. Very often a topic may be introduced by way of a presentational sentence like *I saw a man yesterday*, ... A more explicit formulation of topic chain formation would have to be provided in a more explicit model of discourse grammar.

Osvaldo Jaeggli (personal communication) has suggested to me that the possibility of having an empty topic binding a variable in a given language may be plausibly related to the lack of subject-object ECP effects in such a language. In Chinese, long extraction of a subject (and of an object) appears to be quite free. It has been suggested (Huang (1982)) that such extractions do not show subject-object ECP effects because the subject is always properly governed from within its own clause in this language, perhaps by Infl. Suppose we assume that the Infl, or whatever properly governs the subject, also properly governs the topic position. This would allow the topic to be an empty category, and Chinese would thus be an empty-topic language. This suggestion appears to be extremely plausible, given that Japanese, Korean, and Portuguese also appear to lack standard subject-object ECP effects. (See Zubizarreta (1983) for the lack of these effects in Portuguese. Zubizarreta points out that this fact in Portuguese cannot be explained with the idea of "free inversion" that has been proposed for Italian (Rizzi (1982)). This makes the language look more like Chinese, etc., with true freedom for subject extraction, and makes Jaeggli's suggestion look more plausible, although Zubizarreta argues on other grounds for a different approach, deriving the fact in Portuguese with a strategy analogous to the so-called "*que-qui* rule" in French.) A more thorough discussion of Jaeggli's idea is beyond the scope of this article. (Also see Aoun (forthcoming) for a different approach for deriving the lack of subject-object asymmetries in Chinese.)

grammar, much of this is equivalent to the question why Equi NP Deletion must delete only the subject of a complement clause, but cannot delete an embedded object. The third problem, which concerns zero subject pronouns, is how to account for the fact that some languages allow a zero subject pronoun more freely than others.

I would like to suggest that these two problems may be solved, again not by invoking an ad hoc and new typological parameter, but as the result of the interaction of a number of independently motivated and generalized principles of UG. In particular, I suggest that the relevant facts may be derived jointly by (a) the principle of recoverability, (b) the assumption that a zero pronoun is a pronoun, (c) the assumption that the agreement-marking Agr on a verb qualifies as a potential "antecedent" of a zero pronoun, (d) the binding theory of Chomsky (1981), in particular the condition of disjoint reference (DJR) or condition (B), and (e) the Generalized Control Rule (GCR), extending the ideas of Rosenbaum (1967) and Chomsky (1980; 1981), whose content will be given below.

The principle of recoverability, which says that every empty pronoun must be identified, is a fairly standard assumption, though there is the yet unsolved problem of how it can be made fully explicit. The assumption that a zero pronoun is a pronoun is also in accordance with tradition, as its own widely used name suggests. The assumption that the agreement-marking element in a sentence qualifies as a potential "antecedent" of a zero pronoun (or as an indicator of its content, i.e. reference) follows the lead of Taraldsen (1978) and Chomsky (1981) as a device to capture the "Taraldsen generalization." DJR is indicated in (59); relevant examples illustrating its effect are indicated in (60):[19]

(59) *Disjoint Reference* (DJR)
A pronoun must be free in its governing category.

(60) a. *John$_i$ saw him$_i$.
b. John$_i$ said that Mary saw him$_i$.
c. John$_i$ said that he$_i$ saw Mary.

The GCR is basically Chomsky's (1980) rule of control, extended here to cover both PRO and pro:

(61) *Generalized Control Rule* (GCR)
Coindex an empty pronominal with the closest nominal element.

Roughly, an empty pronominal takes the closest potential antecedent as its antecedent. A nominal element will be understood here to mean either NP or Agr. We will define "closest" in the following manner. Following Chomsky (1980), A is closer to B than C is if A c-commands B but C does not c-command B. Furthermore, for two nodes

[19] Government is defined as in Aoun and Sportiche (1983): α governs β iff α is a lexical category or Agr, and for all φ, φ a maximal category, φ dominates α iff φ also dominates β. According to Chomsky (1981, 211), α is a governing category for β iff α is the minimal category containing β, a governor of β, and a SUBJECT accessible to β. α is accessible to β iff (a) α c-commands β and (b) (potential) coindexation of α and β would not result in the configuration $[_i \ldots i \ldots]$ (the *i*-within-*i* condition).

A and C, both of which c-command B, A is closer to B than C is if A but not C occurs within the same clause as B, or if A is separated from B by fewer clause boundaries than C is. That some notion of "minimal distance" or "local domain" is relevant is a well-known fact about control in English. Thus, in the following sentence the controller of the embedded EC is the closest NP *John,* not the less close *Mary* or the most distant *they*:

(62) They told me that Mary expected that John would promise [*e* to come].

I will depart from Rosenbaum (1967) and Chomsky (1980; 1981) and assume that the notion of "minimal distance" does not distinguish between a c-commanding subject and a c-commanding object within the same clause. The choice between subject and object as the controller of an empty pronominal seems largely determined by pragmatic factors, as has been shown by Manzini (1983), on the basis of examples like (63)–(64):

(63) a. John$_i$ promised Bill [e_i to come].
b. John promised Bill$_i$ [e_i to be allowed to come].
(64) a. John asked Bill$_i$ [e_i to come].
b. John$_i$ asked Bill [e_i to be allowed to come].

2.3. Deriving the Pro-Drop Parameter

We are now ready to derive the facts observed so far. Let us consider sentences of the following forms in "hot," "medium," and "cool" languages in turn:

(65) a. *e* came.
b. John saw *e*.
c. *e* saw *e*.
d. John said that *e* saw Bill.
e. John said that Bill saw *e*.
f. John tried *e* to come.
g. *e* to come.

Consider first "hot" languages. Since English is not a discourse-oriented language, none of the ECs above can be a variable bound to a zero topic. Therefore, each of the ECs, in order to be admitted in English, would have to be a pronominal.[20] Consider (65a) first. According to the GCR, as a zero pronoun, the EC must be coindexed with the closest nominal element. The closest such element in (65a) is the Agr contained in *came*; therefore, the Agr must determine the content of the zero pronoun. But, as is assumed in Chomsky (1982), the Agr in English is too meager to determine its content. Therefore, the sentence violates the principle of recoverability and is thus ill-formed. Consider next (65b), which is also ill-formed, which means that the *e* cannot be a le-

[20] None of the ECs can be anaphors either, since all the NPs in (65) are in thematic position, and an EC is an anaphor only if its antecedent is in nonthematic position.

gitimate zero pronoun. This result can be obtained by a reductio from the GCR and DJR, as follows: Suppose the EC in (65b) is a zero pronoun. Then by the GCR it must be coindexed with a nominal element within its own clause, either the Agr in *saw* or the subject *John*. The Agr is too meager, as before. Therefore, the EC must be identified with the subject *John*. Since a zero pronoun is by assumption a pronoun, it is also subject to DJR, which requires it to be disjoint from *John*. We now derive a contradiction; therefore, the EC cannot be a zero pronoun. Hence the ill-formedness of (65b). The ill-formedness of (65c) needs no further comment. The ill-formedness of (65d) follows in much the same way as that of (65a), since the EC in it must be coindexed with the Agr contained in the embedded *saw*, which is again too meager to determine its content. The EC cannot be coindexed with any of the nominal elements in the matrix clause, since these are not the "closest" nominal elements to it as required by the GCR. In (65e) the EC cannot be identified with any nominal element in the embedded clause, exactly as in (65b). It also cannot be identified with any of the matrix nominal elements, as in (65d). Hence, (65e) is also ill-formed.

We have seen why (65a–e) are all ungrammatical in English. Now consider (65f). The embedded subject EC must be identified with the closest nominal element in accordance with the GCR. There is no Agr in the infinitival clause, so the closest potential antecedent will have to be in the matrix clause. Either the Agr contained in *tried* or the matrix subject *John* may count as the "closest" since they occur within the same clause immediately adjacent to the EC. The meager Agr cannot determine the content of the EC. The subject *John* can, however, and in accordance with the GCR, it must. There is no principle that blocks this mode of identification. Therefore, (65f) is well-formed, with the EC properly interpreted as being controlled by *John*. Finally, (65g) violates the principle of recoverability, since there is no nominal element to identify it. In summary, of all the construction types represented in (65), only sentences having the form of (65f) are grammatical in English.

Let us now turn to the "medium" languages. Since, like the "hot" languages, these are not discourse-oriented languages, none of the ECs in (65) can be a variable bound to a zero topic. Each of them must be a zero pronoun, as before. However, unlike the Agr of English, the Agr of the "medium" languages is rich enough to determine the content of an EC. In Italian or Spanish, the EC in (65a) is properly coindexed with the rich Agr contained in *came*, and the sentence is well-formed. (This mode of coindexation does not contradict DJR, since DJR is concerned with A-binding, but the Agr is not an argument.)[21] The same applies to (65d), where the embedded subject EC is properly coindexed with its own Agr contained in *saw*.

Although (65a) and (65d) are well-formed, sentences of the form (65b), (65c), and (65e) are ill-formed. In (65b) the EC cannot be directly identified by *John* under the

[21] The Agr is not an operator either. Thus, if we more specifically define "Ā-bound" as "operator bound," then the subject EC coindexed with Agr in Italian-type languages is a pronominal by definition, since it is free (in fact, both A-free and Ā-free).

GCR, because of its contradiction with DJR, as before. Although the Agr in *saw* is rich enough to determine the content of the EC, this again will lead to a contradiction. This is because the Agr is subject agreement, and if we identify the object EC with the Agr, we will also necessarily identify it with *John*. The ill-formedness of (65c) and (65e) also follows without further comment.

Next consider sentences of the form (65f), with an EC occurring as the subject of an infinitive. There is no potential antecedent to identify it within the embedded clause, so it must be identified by something in the matrix clause. There are two potential antecedents in the matrix clause, either the matrix subject *John* or the Agr contained in the matrix verb *tried*. Since the Agr agrees with the subject *John*, whether the embedded EC is coindexed with the Agr or directly with the subject *John*, the result is the same: the sentence is well-formed, with the EC bound by *John*. Finally, (65g) violates the principle of recoverability, since there is nothing to determine the content of the EC, and it is thus ill-formed.

In short, in languages of the "medium" type, sentences having the forms (65a), (65d), and (65f) are well-formed, but those having the forms (65b), (65c), (65e), and (65g) are ill-formed.

Let us now consider languages of the "cool" type. Unlike the first two types, the third allows zero topics that are interpreted as discoursally bound to the initial topic of a topic chain. Therefore, each of the ECs in (65) has a potential dual status: it may be a genuine zero pronoun, or it may be a variable that is bound to a zero topic. For each EC in each sentence, then, let us consider both possibilities. In (65a) the EC cannot be a pronominal, because there is nothing to identify it (there being no Agr here). However, it can be a variable bound to a zero topic, since nothing seems to prevent this possibility. Therefore, the sentence is grammatical, with the EC bound by a zero topic, i.e. referring to a discourse topic. In (65b) the EC also cannot be a pronominal, because of the interaction of the GCR and DJR, but again it can be a variable. Thus, the sentence is again grammatical, with the EC referring to a discourse topic. In (65c) both ECs are again not pronominals, but each again can be a variable. The view that both ECs can be variables requires one to assume that a sentence can have multiple topics. But precisely this assumption appears to be necessary, given well-formed sentences like (66):[22]

(66) Zhangsan, neiben shu, ta hen xihuan.
Zhangsan that book he very like
'(As for Zhangsan), that book, he likes very much.'

Skipping (65d) for one moment, we can easily see that the same applies to (65e). The EC in (65e) cannot be a pronominal, because of the GCR and DJR, but it can be a variable. Thus, the sentence is grammatical, with the EC unambiguously interpreted as bound by a zero topic.

[22] It is not important whether each of the two NPs preceding the subject is literally identified as a topic. What is clear is that each of these NPs is in an operator position, a position adjoined to some clausal node, so that any empty category bound to such a position is a variable.

Returning to (65d), we see that the embedded subject EC can be either a pronominal or a variable. It can be a pronominal, since it can be properly bound by *John,* the closest nominal element c-commanding it (there being no Agr). It also can be a variable bound by a zero topic, since nothing seems to prevent this. The sentence is therefore grammatical and ambiguous, with the EC referring to *John* (as a pronominal) or to a discourse topic distinct from *John* (as a variable). Note that the EC in (65d) in Chinese, when a pronominal, differs from its counterpart in Italian or Spanish in that, in the former case, the EC is identified by way of control by an NP (on a par with a controlled PRO in an infinitival clause), and in the latter, the EC is coindexed with Agr.

Consider now the status of (65f) with a subject EC in an infinitival clause. It is already clear that the EC can be a pronominal, since it can be identified by *John,* as in (65d). However, unlike (65d), the EC here *must* refer to *John*, as the following example shows:

(67) Zhangsan$_i$ shefa e_i bangmang wo.
Zhangsan try help I
'Zhangsan tried to help me.'

(68) *neige ren$_i$, Zhangsan shefa e_i bangmang wo.
that man Zhangsan try help I

This means that the EC in (65f) cannot be a variable. Although nothing said so far gives this result, it is plausible to assume that it can be derived from an independent principle, namely, the Empty Category Principle (ECP) of Chomsky (1981), which requires that every nonpronominal EC be properly governed.[23] The assumption that the embedded subject position in (67) is not properly governed not only enables us to rule out a variable in that position, but also accounts for the fact that the same position admits no lexical NP:[24]

(69) *Zhangsan shefa Lisi bangmang wo.
Zhangsan try Lisi help I

[23] α properly governs β iff α governs β and either (a) α is a lexical category (but not Agr) or (b) β is coindexed with α.

[24] There is a technical problem concerning the application of the government theory in languages like Chinese. The subject of what is intuitively a finite clause must be considered governed since it can be lexically realized and Case-marked in accordance with Case theory, though the Case marker of the subject cannot be Agr, as assumed in Chomsky (1981), there being no Agr in the language. On the other hand, the subject of what is intuitively a nonfinite clause must be considered ungoverned, since it cannot be lexically realized. This is the way we would want the government theory to work. The technical difficulty here is that there is no systematic overt marking for finiteness in Chinese. It is generally agreed that there is no tense marking in the language. It seems that finiteness is not identified with the presence of tense in all languages. In Chinese, what is intuitively a finite clause may be distinguished from a nonfinite clause by the former's but not the latter's *potential* possibility of occurring with a marking for an aspect, say the perfective aspect *le* or a variant of it. (For some discussion of what counts as a finite clause in a given language, see George and Kornfilt (1981).) I will assume that there is an Infl node (which may be phonetically empty) in finite clauses in Chinese that governs the subject.

Finally, the EC in (65g) is neither a pronominal nor a variable, and the sentence may be considered ungrammatical in Chinese also.[25]

In languages like Chinese, then, sentences of the forms (65a–f) but not (65g) are grammatical. Those having the form (65d) are furthermore ambiguous, the EC being either a pronominal or a variable. Those having the other forms, however, are unambiguous. The ECs in (65a–c) and (65e) are variables, whereas the EC in (65f) is a pronominal.[26]

Summarizing, I have suggested that the problems posed in section 1 may be answered as follows. First, the fact that some languages allow zero topics binding variables whereas others do not is derived as a special case of the "discourse-oriented vs. sentence-oriented" parameter. Second, none of our languages allows a genuine zero object pronoun, because of the interaction of the GCR and DJR. Finally, the distribution of a zero subject pronoun is closely tied to the presence or absence of a potential antecedent rich enough in content (Agr or an actual NP). In particular, the phenomenon of subject "pro drop" can occur when there is a rich agreement element, or when there is no agreement at all. The first case occurs in languages like Italian and Spanish, where there is an Agr rich enough to identify the content of a zero subject. The second case occurs in languages like Chinese and Japanese, where there is no Agr at all, in which case a zero subject pronoun is identified by an NP in a superordinate clause. Note that the second case also occurs in *all* three types of languages when the zero subject pronoun occurs in an infinitival clause. This is the standard case of controlled PRO. Therefore, the distribution of a zero subject in a tensed sentence of "cool" languages is on a par with that of a zero subject in a tenseless clause of *all* languages. The only case where a zero subject cannot occur is where there is an agreement element that is not rich enough in content, as in English. The mere presence of an Agr in English requires it to be coindexed with its zero subject, but its degenerate nature prevents it from fulfilling the condition of recoverability.

3. Further Consequences

In the preceding pages I have argued for a way to sort out the facts considered in section 1 and have suggested a line of research that points to a possible solution to the problems posed. The suggestion seems to me to be highly plausible, and it has certain additional desirable consequences. In this section I will show that it provides a possible solution to certain otherwise peculiar differences between English and Chinese in three types of

[25] We may make this assumption on the grounds that there is no independent sentence in the language whose subject has a thematic role but must *not* be lexically realized. (Given that even imperatives allow their subjects to be lexicalized, we may assume that the imperative is a finite sentence and its subject is governed.)

[26] Note that Portuguese differs from Chinese in that each EC identified as a pronominal in (65) is so identified on a par with such an EC in Italian or Spanish. In German, only the EC in (65g) may be identified as a pronominal, and *at most* one EC in the other sentences may be identified as a variable, because of its requirement that only one NP may appear before the main verb.

syntactic constructions. These are (a) constructions involving the "strong crossover" phenomenon, (b) constructions related to Subjacency, in particular Ross's Complex NP Constraint (CNPC), and (c) constructions related to Ross's Left Branch Condition (LBC). In each case I will show that there is an asymmetry in Chinese that does not exist in English. The typological differences can be plausibly reduced in each case to the differences already observed between the two languages.

3.1. Strong Crossover

The following sentences illustrate the phenomenon of "strong crossover" (Postal (1971), Wasow (1972)):

(70) a. *John$_i$, he$_i$ said e_i saw Bill.
b. *John$_i$, he$_i$ said Bill saw e_i.
(71) a. *Who$_i$ did he$_i$ say e_i saw Bill?
b. *Who$_i$ did he$_i$ say Bill saw e_i?
(72) a. *This is the man who$_i$ he$_i$ said e_i saw Bill.
b. *This is the man who$_i$ he$_i$ said Bill saw e_i.

In each of these sentences the empty category *e* is coindexed both with a c-commanding subject (i.e. *he*) and with a nonargument (the topic in (70) and the Comp in (71) and (72)). According to traditional analyses, each of these sentences involves the movement of an NP in the position marked by *e* across a coreferential c-commanding NP into a Comp or topic position (cf. Chomsky (1977)). In each of the (a) sentences, movement originates from a subject position, and in each of the (b) sentences, it originates from an object position. Since both the (a) and the (b) sentences are ungrammatical, we know that strong crossover is prohibited in English, whether movement takes place from a subject or an object position. Now consider constructions having the same forms in Chinese:

(73) a. Zhangsan$_i$, ta$_i$ shuo e_i mei kanjian Lisi.
Zhangsan he say no see Lisi
'Zhangsan$_i$, he$_i$ said that [he$_i$] didn't see Lisi.'
b. *Zhangsan$_i$, ta$_i$ shuo Lisi mei kanjian e_i.
Zhangsan he say Lisi no see
'Zhangsan$_i$, he$_i$ said that Lisi didn't see [him$_i$].'

What is interesting here is the asymmetry between (73a) and (73b). These two sentences have exactly the same form as their English counterparts in (70). However, whereas (70a) and (70b) are both ill-formed in English, only (73b), but not (73a), is ill-formed in Chinese. In other words, in Chinese, strong crossover involving an embedded object EC is prohibited as in English, but strong crossover involving an embedded subject

EC is permitted, contrary to the cases in English.[27] A similar asymmetry can be observed under relativization. To see this asymmetry, however, we must first look at the following data:

(74) a. neige ren$_i$, ta$_i$ hen lei.
that man he very tired
'That man, he was very tired.'
b. *[ta$_i$ hen lei] de neige ren$_i$
he very tired DE that man
'the man who was very tired'
c. [e_i hen lei] de neige ren$_i$
very tired DE that man
'the man who was very tired'

(74a) shows that in a topic-comment sentence, the NP in the comment that is co-referential with the topic may be a pronoun when it appears as the subject of the comment clause. (74b–c) show, on the other hand, that in a relativized construction, if the subject of the relative clause is coindexed with the head, it must be an EC. Making the necessary adjustments, then, the Chinese counterparts of the English relativized NPs in (72) are as follows:

(75) a. [e'_i shuo e_i mei kanjian Lisi] de neige ren$_i$
say no see Lisi DE that man
'*the man who$_i$ [he$_i$] said e_i didn't see Lisi'
b. *[e'_i shuo Lisi mei kanjian e_i] de neige ren$_i$
say Lisi no see DE that man
'*the man who$_i$ [he$_i$] said Lisi didn't see e_i'

In both (75a) and (75b) e' corresponds to the pronoun *he* in the English sentences in (72), and e corresponds to the EC in (72). The difference between (75a) and (75b)

[27] There is a clear connection between the sentences (73a–b) and (22a–b) discussed earlier. The asymmetry shown in (22a–b) may be seen to directly give rise to the asymmetry in (73a–b). For discussion of the existence of strong crossover effects in Japanese and its implications, see Saito (1982). Against the unacceptable sentence (73b), an LI reviewer has suggested the following as a possible counterexample:

(i) Li Xiaojie$_i$, ta$_i$ shuo [wo meiyou qing e_i, suoyi buken lai].
Li Miss she say I not invite so unwilling come
'Miss Li, she said I didn't invite her, so she was unwilling to come.'

In this sentence, note that there is an empty subject in the second conjunct of the coordinate embedded clause immediately preceding 'unwilling to come'. The embedded object EC in the first conjunct thus need not directly depend upon the matrix subject 'she' for its reference, but may derive its reference from the non-c-commanding empty subject following it. Such a structural configuration is on a par with sentences like (100)–(102) discussed later and may be dealt with along lines suggested for those sentences. It is interesting to note that without the second conjunct of the embedded coordinate clause, the object EC following 'invite' does not normally refer to 'Miss Li':

(ii) *Li Xiaojie$_i$, ta$_i$ shuo [wo meiyou qing e_i].
Li Miss she say I not invite

relevant to our discussion is in the position of e (not e'). In (75a) e occurs in embedded subject position, and in (75b) it occurs as an embedded object. Thus, the structural difference between (75a) and (75b) in Chinese is on a par with that between (72a) and (72b) in English. Again we see a subject-object asymmetry in Chinese that does not exist in English.

The systematic difference between Chinese and English under strong crossover can be considered a direct consequence of the fact that Chinese is a pro-drop language allowing a zero subject pronoun, whereas English is not. In particular, consider how both the (a) and (b) sentences in (70)–(72) in English can be ruled out. According to Chomsky's (1981) functional definition indicated in (37), each EC is locally identified as a pronominal or a nonpronominal. In all of the sentences in (70)–(72), the EC is *locally* bound by *he*, which is in a thematic position. By definition, then, each of these ECs cannot be a nonpronominal. However, none of these ECs can be pronominals either. A zero object pronoun in each of the (b) sentences is excluded by the GCR plus DJR, and a zero subject in each of the (a) sentences is excluded by the requirement that it be coindexed with the Agr of its verb and the fact that the Agr is too meager. Since none of these ECs can be either a pronominal or a nonpronominal, all of (70)–(72) are ill-formed.

Consider now the sentences in Chinese. By definition (37), each of the ECs in (73) is a pronominal. However, the object EC in (73b) cannot be a pronominal, for reasons already seen, so (73b) is ill-formed. On the other hand, the subject EC in (73a) can be a legitimate pronominal since it is identified by the closest c-commanding NP, its local binder *ta* 'he'. Therefore, the sentence is grammatical. In other words, regardless of its derivational history, sentence (73a) emerges as equivalent to a well-formed structure of left dislocation. The asymmetry between (73a) and (73b) thus follows. It should be easy to see that the asymmetry between (75a) and (75b) falls out in the same way.

3.2. Subjacency

There is also an interesting difference between Chinese-type languages and English-type languages with respect to Subjacency, in particular the CNPC. As is well known, violation of the CNPC in English gives rise to ungrammatical strings, whether the complex NP occurs in subject or in object position:

(76) a. *John$_i$, the voice with which e_i sings is good.
b. *John$_i$, I like the voice with which e_i sings.

(77) a. *John$_i$, the books which e_i wrote are many.
b. *John$_i$, I have read many books which e_i wrote.

If we turn to the Chinese counterparts, however, an important asymmetry emerges:

(78) a. Zhangsan$_i$, e_i changge de shengyin hen haoting.
Zhangsan sing DE voice very good-to-hear
'Zhangsan$_i$, the voice with which [he$_i$] sings is good.'

b. *Zhangsan$_i$, wo hen xihuan e_i changge de shengyin.
Zhangsan I very like sing DE voice
'Zhangsan$_i$, I like the voice with which [he$_i$] sings.'

(79) a. Zhangsan$_i$, e_i xie de shu bu shao.
Zhangsan write DE book not few
'Zhangsan$_i$, the books that [he$_i$] wrote are not few.'

b. *Zhangsan$_i$, wo nian-le bu shao e_i xie de shu.
Zhangsan I read-LE not few write DE book
'Zhangsan$_i$, I have read quite a few books that [he$_i$] wrote.'

The asymmetry is again a subject-object asymmetry. In (78a) and (79a) "extraction" is permitted from a complex NP in subject position, but in (78b) and (79b) such an extraction is impossible from a complex object NP. The grammatical (a) sentences constitute a violation of Subjacency under the assumption that the sentence is derived by movement. This violation, however, cannot be taken as evidence that Subjacency does not apply to Chinese. As should be obvious, taking this stand does not explain the asymmetry under consideration, and would wrongly admit the ill-formed (b) sentences.

I suggest that the difference between Chinese and English can again be derived largely from what we have assumed on independent grounds. First, it is fairly well known that the topic position of a sentence must be available at D-structure (given left-dislocated sentences and sentences of the form (56)). Therefore, for each variable bound to a topic, there are two possible ways to derive it: it may be created by movement as a *wh*-trace, or it may start out as an EC at D-structure and later be coindexed with the topic (and become a variable). Let us then make this null hypothesis and consider each of these sentences. In (76a) and (77a) movement from the position of the EC to that of *John* is excluded by Subjacency. In order to be admitted, then, *e* would have to be base-generated, and before coindexing takes place, it is by the functional definition (37) a pronominal. As a pronominal, it must be identified by the meager Agr of *sings,* but not by the topic. Thus, none of the subject ECs can be derived either by movement or by nonmovement. The same applies to the object ECs in (76b) and (77b). Therefore, all of (76) and (77) are ill-formed.

Consider now the Chinese sentences in (78) and (79). Assuming that Subjacency also obtains in Chinese, none of the ECs here can be derived by movement, as before. The EC must therefore be base-generated. In (78a) the EC as a pronominal must be identified by the closest potential antecedent. There is no nominal element within the relative clause containing *e* in (78a). The head of the relative clause is the closest nominal element, but clearly it cannot serve as the antecedent for the EC, since it is already the antecedent of something else (the relativized instrument of 'sing'). Suppose we stipulate that the head of a relativized construction that is already coindexed with something does not count as a potential antecedent of another EC.[28] This will enable us to look further

[28] Presumably this may be a consequence of a properly formulated version of the *i*-within-*i* condition on accessibility (see footnote 19).

up. The next closest nominal element is the topic *Zhangsan*, which may be coindexed with the EC. (78a) is thus well-formed because the EC may start out as a pronominal and become a variable when it is coindexed with *Zhangsan* in accordance with the GCR. The EC in (78b), on the other hand, cannot be derived in the same manner. This is because the closest potential antecedent for the EC in (78b) is the matrix subject *wo* 'I', not the topic *Zhangsan*. Therefore, the sentence is ill-formed.[29] The same applies to (79).[30]

Incidentally, although extraction from a complex object NP is not allowed, such extraction is possible once the complex NP itself is topicalized. Thus, compare the ill-formed (78b) with the well-formed (80):

(80) Zhangsan$_i$, e_i changge de shengyin wo hen xihuan.
Zhangsan sing DE voice I very like
'Zhangsan$_i$, the voice with which [he$_i$] sings, I like.'

This contrast shows that the application of the GCR is structure-dependent, and not dependent on grammatical relations like subject, object, etc.[31]

[29] For some reason, actually, the EC in (78b) does not refer to the subject 'I' either, but is more likely to be interpreted as having arbitrary reference, so that the comment clause in (78b) alone would mean something like 'I like the voice of one's singing'. This shows that the GCR as stated in (61) needs some refinement to allow for cases of arbitrary pronominal ECs.

[30] Consider the following sentences:

(i) Zhangsan$_i$, ni zhidao e_i changge de shengyin hen haoting.
Zhangsan you know sing DE voice very good-to-hear
'Zhangsan$_i$, you know the voice with which [he$_i$] sings is good.'
(ii) *Zhangsan$_i$, ni zhidao wo hen xihuan e_i changge de shengyin.
Zhangsan you know I very like sing DE voice
'Zhangsan$_i$, you know I like the voice with which [he$_i$] sings.'

The contrast between (i) and (ii) is on a par with that between (78a) and (78b). The difference is that the comment clause that contains the complex NP in question in (78) is now embedded under 'you know' in (i) and (ii). According to formulation (61) of the GCR, the EC in (i) cannot be directly coindexed with the topic *Zhangsan* any more than the EC in (ii). However, one may assume that *Zhangsan* in (i) originates from the operator position of the complement clause under 'you know'. After the GCR applies to coindex the EC with *Zhangsan* in the embedded operator position, the latter undergoes movement into the matrix topic (or Comp) position.

[31] Chinese appears to lack the effects of Ross's Sentential Subject Constraint (SSC), as indicated in (i):

(i) Zhangsan$_i$, [[e_i kan zhebu dianying] bu heshi].
Zhangsan see this movie not appropriate
'Zhangsan$_i$, that [he$_i$] sees this movie is not appropriate.'

If the SSC is taken to be a subcase of Subjacency, then this lack of SSC effects in Chinese, but not in English, can be accounted for in the same way as (78a) and (78b). In Huang (1982), however, I assume that the SSC is to be derived from the theory of proper government, as a subcase of the Condition on Extraction of Domain (CED), which requires that every domain out of which extraction takes place must be properly governed. Since the subject position in Chinese finite clauses is properly governed (see the discussion in footnote 18), the fact that (i) is well-formed in seeming violation of the CED is expected, independent of what is said about (78a) and (78b). (See Kayne (1981, 1983) and Pesetsky (1982), where the SSC is also considered unrelated to Subjacency.)

3.3. The Left Branch Condition

The following sentences indicate that extraction of the possessive NP in violation of Ross's LBC is impossible in English whether the NP headed by *father* is in subject or in object position:

(81) a. *Whose$_i$ was [e_i father] very rich?
b. *Whose$_i$ did you see [e_i father]?
(82) a. *John$_i$, [e_i father] was very rich.
b. *John$_i$, I saw [e_i father].

However, there is again a subject-object asymmetry in Chinese:[32]

(83) a. Zhangsan$_i$, [e_i baba] hen youqian.
Zhangsan father very rich
'Zhangsan, [his] father is very rich.'
b. *Zhangsan$_i$, wo kanjian [e_i baba] le.
Zhangsan I see father LE
'Zhangsan, I saw [his] father.'
(84) a. [$_{NP}$[$_{S'}$[e_i baba] hen youqian de] neige xuesheng$_i$] lai le.
father very rich DE that student come LE
'The student whose father is rich came.'
b. *[$_{NP}$[$_{S'}$ wo kanjian [e_i baba] de] neige xuesheng$_i$] lai le.
I see father DE that student come LE
'The student whose father I saw came.'

The asymmetry in (83) and (84), and the lack of it in (81) and (82), can again be plausibly derived in a similar way. In particular, each of the (a) sentences in Chinese may be derived by generating an empty pronominal that becomes a variable after it is properly identified with the topic or the head of a relative, and the (b) sentences may be excluded because the pronominal cannot be properly turned into a variable. On the other hand, all the English sentences may again be excluded by assuming that none of the ECs in them may be a pronominal.[33]

[32] I assume that inalienable possession nouns differ from other nouns in that they obligatorily assign a thematic role Possessor to an argument, whereas other nouns need not do so. Therefore, the assumption that there is an EC in each of (81)–(84) is a consequence of the thematic theory and the Projection Principle.

[33] One might assume that there is a meager Agr on every head noun, given evidence of determiner-head agreement like *these books, *these book,* etc.

In this connection, consider the following sentence, which appears to contradict what is said about the ill-formedness of (83b) and (84b):

(i) juzi, wo buo-le pi le.
orange I take-off-LE skin LE
'The orange, I have peeled (the skin of).'

If we assume that the topic 'orange' binds a possessive EC immediately preceding the object 'skin', what we have said about (83) and (84) will rule out (i) as ill-formed. There is one possible way to admit this sentence, however. Chinese exhibits numerous examples of compound formation by which a verb-object phrase is "com-

4. Conclusion

In this article I have argued that the problem of formulating an appropriate parameter to account for the distribution and reference of what is usually called a "zero pronoun" in various languages can be best approached by recognizing that there are two distinct parameters involved. One parameter distinguishes zero-topic from non-zero-topic languages, and the other distinguishes pro-drop from non-pro-drop languages. The possibility of allowing a variable bound to a zero topic is reduced to a more general parameter that distinguishes discourse-oriented from sentence-oriented languages. The Pro-Drop Parameter is derived in a highly modular manner, by a number of independently motivated principles of grammar, which also provide a plausible account for a number of asymmetries that exist in Chinese but not in English.

If the approach outlined here is on the right track, it is worthwhile to point out a number of nontrivial theoretical implications.

Recall first that the GCR has been stated as a rule of coindexing. Suppose that each noun phrase that has not been assigned an index by the S-structure level is assigned one at this level (Chomsky (1981)). Then the GCR can be equivalently understood as a condition on the application of index assignment. A question that may arise now is

---

pressed" into a lexical category, e.g. *dan-xin* 'hold-heart' in *wo dan-xin ta* 'I worried about him'. The sequence *buo* (*-le*) *pi* 'take skin off' may be considered a product of this perhaps not yet completed process of compound formation, so that it may now be regarded as a semicompound ("semi-" because the marker *-le* may intervene between the two components) meaning 'peel'. If so, then there is no possessive EC immediately preceding *pi* 'skin' in (i), and the topic binds an object EC immediately *following* the transitive semicompound 'take-off-skin'. The question that now remains is why one cannot say the same thing about the sequence 'see-father' in each of the ill-formed sentences in (83) and (84). I think that the answer will have to come from a more articulate theory of reanalysis and of the notion of what may count as a "possible word" (cf. Hornstein and Weinberg (1981)). The notion of what counts as a "possible" transitive verb-compound seems to be closely related to the notion of basic transitivity. 'Take the skin off' may be taken as a verb expressing an action that directly or indirectly affects something or someone, so it may be used as a transitive verb. 'See the father', on the other hand, can hardly affect someone, and is therefore not used as a transitive semicompound. It is interesting to note that the following sentence, in which 'kill the father' may be more easily understood as expressing an action that affects someone, is considerably more acceptable than (83b) and (84b):

(ii) ?Zhangsan, tufei dasi baba le.
Zhangsan bandit kill father LE
'Zhangsan, the bandits killed his father.'
(Lit.: Zhangsan, the bandits "father-killed" him.)

Compare also the following sentences, where a preverbal *ba*-object occurs in overt form. (The presence of a *ba*-object is taken by Thompson (1973) as a sign of transitivity par excellence in Chinese.)

(iii) wo ba juzi buo-le pi le.
I BA orange take-off-LE skin LE
'I peeled the orange.'

(iv) tufei ba Zhangsan dasi-le baba.
bandit BA Zhangsan kill-LE father
'The bandits killed Zhangsan's father.'
(Lit.: The bandits "father-killed" Zhangsan.)

(v) *wo ba Zhangsan kanjian-le baba.
I BA Zhangsan see-LE father
'I saw Zhangsan's father.'
(Lit.: I "father-saw" Zhangsan.)

whether it is possible to construe the GCR as a condition on representation, in effect an "extended" binding condition governing empty pronominals, thus allowing the statement of a maximally simple rule of free indexing. Such a binding condition would require, say, that a pronominal EC must be bound by the closest nominal element. It should be pointed out that this conception of the GCR would be undesirable. The reason involves our adoption of the view that the "allo-category" (or allomorphy) of an EC is determined locally by principle (37) and that the identity of an EC may change in the course of a derivation. In the discussion of the subject-object asymmetry with respect to extraction from a complex NP in Chinese, it was indicated that a variable located within a complex NP may start out as a pronominal, be coindexed with a topic in accordance with the provisions of the GCR, and end up as a variable by definition. Whether a given complex NP occurs in subject or in object position thus makes a difference in whether a given pronominal EC can be properly converted into a variable. If the GCR were to be reformulated as a condition on the output of free indexing, however, the subject-object asymmetry described above would be left unaccounted for. As a result of free indexing, an EC that starts out as a pronominal may be coindexed with the topic whether the complex NP containing it is in subject position (as in (78a) and (79a)) or object position (as in (78b) and (79b)). After coindexing takes place, however, the EC becomes a variable and is not subject to the GCR. One would then wrongly predict that both the (a) and (b) sentences of (78) and (79) are well-formed. Note that without any additional assumptions, such a conception would also wrongly admit English sentences like (85), in which the EC could start out as a pronominal and be freely converted into a variable:[34]

(85) *Which event$_i$ did you read stories that describe e_i?

If the GCR should be construed as a rule (or equivalently as a condition on index assignment), rather than as a condition on the output of free indexing, note that there is evidence that the rule Move α must be an available option in UG. This is, plainly, because there are well-formed sentences containing variables that could not have been derived by the GCR. In English, all well-formed sentences containing variables must be generated by movement, since no variable may originate as a pronominal under the provisions of the GCR. In Chinese, although some variables may be derived from pronominals by the GCR, movement must also be available because typical topicalized sentences like the following must be admitted, where each EC is locally bound by some-

[34] In Chomsky (1982) the possibility of deriving the variable in (85) from a base-generated pronominal is excluded by the stipulation that indexing involves only A-positions in syntax, and $\bar{A}$-indexing does not happen until after a derivation enters LF. In order to be bound by the topic in LF, the EC in (85) would therefore be free (A-free and $\bar{A}$-free) at S-structure. Since it is governed, it cannot be a PRO at this level. Since it is unidentified, it also cannot be pro. Hence the ill-formedness. Note, however, that the same result can be obtained without this stipulation, if the GCR is construed as a condition on rules. Furthermore, note that even if one accepts the stipulation that $\bar{A}$-indexing applies only in LF, this point still holds. The asymmetry between (78a) and (78b) in Chinese shows that the application, not the output, of $\bar{A}$-indexing must be constrained by the provisions of the GCR.

thing other than the closest nominal element:

(86) neige ren$_i$, Zhangsan kanjian e_i le.
that man Zhangsan see LE
'That man, Zhangsan has seen.'

(87) neige ren$_i$, Zhangsan shuo Lisi kanjian e_i le.
that man Zhangsan say Lisi see LE
'That man, Zhangsan said Lisi has seen.'

This point is worth mention because the opposite has been suggested by many (for example, Bresnan (1982)), based upon the alleged restrictiveness and simplicity that one can achieve by eliminating Move α in favor of a simple rule coindexing base-generated ECs. One could reject this view simply by the fact that assuming that Move α may or may not be involved in each case (that it exists as an optional rule) is the null hypothesis. But there is empirical evidence also that the two processes cannot be collapsed as one, since a base-generated EC is subject to the GCR but a moved trace is not. In Chinese, it has been suggested, based upon sentences like (56) for which there are no nontopicalized sources, that movement is never involved in the generation of any topic-comment sentence in this language (Li and Thompson (1976b)). Such an assumption again can be dismissed by the argument that it is not the null hypothesis (there being no independent evidence to rule out movement as a possibility) and by the theory-internal evidence derived from (86)–(87).

The existence of some process (movement) distinct from the process of coindexing a base-generated EC can also be derived from the fact that Subjacency must be construed as a condition on movement, rather than as one on coindexing or on output representations. This is, again, apparent from the sentences in (76)–(79). The well-formedness of (78a) and (79a) shows that neither the process of coindexing under the GCR, nor the output of coindexation, is subject to Subjacency. On the other hand, the ill-formedness of all of (76) and (77) and the (b) sentences of (78) and (79), together with the well-formedness of (86) and (87), shows that there is a distinct process (movement) whose application is subject to Subjacency.

Incidentally, a similar observation can be made about the LBC, or whatever principle it is a special case of. Given the asymmetry observed in Chinese, it is clear that the LBC cannot be formulated as a condition on representations. This asymmetry argues against the account proposed by, say, Gazdar (1981) under his Generalized LBC.

Note that both the GCR and Subjacency are most naturally considered to apply in the syntax by S-structure. In the case of the GCR, this is because index assignment must apply by this level so that binding theory may apply at this level.[35] Recall also that,

[35] See Chomsky (1981, 196ff.) for evidence that the binding theory must apply (at least) at S-structure. Consideration of strong crossover sentences like *Whose mother did he see?* suggests that the binding theory must also apply at LF (after the sentence has been reconstructed as "For which x, he saw x's mother" (Chomsky (1976))). See Aoun (1982) for additional arguments that the binding theory obtains also in LF.

since we have derived the absence of a zero object pronoun by a reductio of the GCR and DJR of binding theory, it is also necessary that the effect of the GCR be available at some level where binding theory applies. As for Subjacency, there is sufficient evidence that it has effect only in the syntactic, but not the LF, module of grammar, assuming that LF includes rules that displace quantificational NPs and *wh*-phrases in situ.[36] Since both the GCR and Subjacency apply in the same module of grammar, and they do not apply to the same process, it is clear that there are two syntactic processes: movement and coindexing.

I will conclude this article by indicating that the line of research outlined above is not without problems whose full solution must await further research. I will indicate some of these problems and speculate on what might be a plausible answer to each of them. I believe, however, that these are problems of execution, and that the main points of the article may remain largely unaffected however the details are worked out.

One problem is the vagueness that has characterized any attempt to relate pro drop to the presence of a rich agreement system. Regardless of the plausibility of the theory as suggested by the evidence from Pashto, there is the question of "how rich is rich enough" for a language to allow pro drop. German and French, for example, have fairly rich systems of agreement, but they have both been described as non-pro-drop languages. It may be that the presence of a rich agreement system only makes a language a potential pro-drop language. Something else may be required in each language to make pro drop possible. It is also possible that one must learn, from positive evidence or from indirect negative evidence, whether a potential pro-drop language actually allows pro drop.[37]

Another problem concerns the well-known fact that certain empty pronominals may be arbitrary in reference and need not be bound. Examples include the following:

(88) It is important *e* to abide by one's own words.
(89) *e* smoking is harmful.
(90) John wondered how *e* to sell oneself.

Our reliance on the notion of recoverability and the formulation of the GCR do not permit these sentences to be acceptable. It is apparent that the condition of recoverability must be relaxed for a free pronominal EC in certain positions. A plausible characterization of these positions has been given in Manzini's (1983) treatment of ungoverned

[36] See Huang (1982) for arguments that Subjacency does not, though the ECP does, obtain in the LF module. It is possible to adopt the idea of "pied-piping" in LF so that the applicability of Subjacency may be preserved here (along plausible lines suggested in Nishigauchi (1983) and similar lines in Kayne (1983)), but I believe this has undesirable consequences. To show the latter point here, however, would take us too far afield.

[37] See Jelinek (1983), who cites evidence from Arabic and suggests that the crucial element that makes Agr rich enough is person marking. Jim Higginbotham has pointed out correctly to me that it is misleading to talk about a rich Agr being capable of identifying or recovering in full the content of an EC. Rather, the Agr can only determine the EC's content, but cannot specify its reference uniquely. Higginbotham suggests that the possibility of pro drop in a richly inflected language is probably best thought of as a blind feature of formal grammar, not semantically based. If anything is identified by Agr, it is the formal features of the missing subject, not its reference.

PRO using the notion "domain-governing category."[38] One might try to incorporate and somehow extend her idea, though the exact details remain to be worked out, since her theory concerns only ungoverned PROs, whereas we have been dealing here with other ECs.

Still another problem is the apparent partial redundancy between the GCR and binding theory. In particular, the absence of a pronominal EC in object position is accounted for by DJR and the GCR, and also partially by the first two conditions of binding theory, that is, DJR and the condition that an anaphor is bound in its governing category. The two binding conditions require all pronominal anaphors (PROs) to be ungoverned, thus excluding them from object position. However, since certain pronominal ECs (pro's) can be governed, the redundancy observed here is only partial. It also does not seem possible to entirely collapse the generalized control theory outlined here with binding theory. Binding theory refers to A-binders only, but control theory refers to closest potential binders that may be in an A- or Ā-position. Furthermore, control theory is construed as a rule or a condition on rules, though binding theory is generally assumed to be a set of conditions on representation.

Finally, our discussion of empty pronominals will not be complete without mention of some constructions that may appear to be counterexamples to the claim that none of our languages allows object pronominal ECs.[39] These include sentences like the following, which have been treated in Chomsky (1980):

(91) John bought a book$_i$ for Mary to read e_i.

(91) contains an object EC that is apparently locally bound to a thematic position. Similar examples in Chinese can be easily constructed:

(92) Zhangsan mai le yiben shu$_i$ gei Lisi kan e_i.
Zhangsan buy LE one book for Lisi read
'Zhangsan bought a book for Lisi to read.'

Since it is otherwise clear that an object EC cannot be a pronominal, the existence of such examples as these should lead one to suspect that they are not genuine counterexamples. Chomsky (1980) proposes that the EC in (91) is in fact a variable bound by an empty operator in the embedded Comp, which is coindexed with 'a book' by a rule of "predication." We might assume that this rule is a special case of the GCR. It is

[38] According to Manzini, α is a domain-governing category for β iff (i) α is the minimal c-domain of a subject that contains β and a governor of the c-domain of β, and (ii) α contains a SUBJECT accessible to β. Roughly, a PRO in an obligatory control structure has a domain-governing category by this definition, but in sentences like (88)–(90) the *e* has no domain-governing category (or governing category) and thus need not be bound.

[39] The question whether this claim can be maintained as a universal cannot be fully dealt with here. Cole (1982) and Mohanan (1983) appear to pose some problems, but my lack of knowledge of the languages they deal with prevents me from discussing them in any detail.

apparent that the same assumption, if correct, can be extended to the Chinese counterpart (92). A question that arises here is how to ensure that this assumption is available only in connection with sentences like (91)–(92), but not in connection with the type of examples used to show our point earlier, or with ungrammatical sentences like (93):

(93) *John said Bill would see *e*.

In other words, why can't the EC be a variable bound by an empty operator in the embedded Comp that is coindexed with *John* by the rule of predication? It is plausible to assume that the relevant difference between (91)–(92) and (93) is that the embedded clause in (93) is an argument but the purposive clause in (91)–(92) may be taken as a modifier or a predicate. In other words, (91)–(92) involve structures of predication, whereas (93) involves a structure of complementation. A modifier may be used to say something "about" some NP, but a complement is not used to make a comment "about" some NP. If this is the relevant distinction here, one may assume that the rule of predication (or the GCR), when it coindexes the operator position of a given clause with an external argument position, is applicable only when the clause is a structure of predication, and not if it is a complement clause. That this is probably on the right track is suggested by the fact that other examples that appear to contain object pronominal ECs all seem to involve structures of predication, rather than structures of complementation. For example, in addition to the purposive constructions above, sentences of the following form in Chinese are very common, in which each embedded object EC is coindexed with the matrix object:

(94) Zhangsan you yiben shu$_i$ [wo kan-bu-dong e_i].
Zhangsan have one book I don't-understand
'Zhangsan has a book, which I don't understand.'

(95) Zhangsan mai le yidong fangzi$_i$ [wo hao xihuan e_i].
Zhangsan buy LE one house I very like
'Zhangsan bought a house, which I really like.'

As indicated in the English translations, the strings *wo kan-bu-dong* and *wo hao xihuan* both have roughly the semantic status of a nonrestrictive relative clause. It is generally agreed, however, that they do not have the syntactic status of a relative clause or form a constituent with the object, since relative clauses (restrictive or otherwise) must precede their heads. In Li and Thompson (1981, 611–620), these clauses and purposive clauses like the one in (92) are called *descriptive clauses*. It should be clear also that each of these clauses does not serve as the complement of the verb 'have' or 'bought'. Rather, each seems to be a clause that predicates on the matrix object. The EC in each of (94) and (95) thus again may be analyzed, as in the case of the purposives, as a variable bound by an abstract operator in the embedded Comp coindexed with the matrix object.

This analysis, if correct, may be extended to offer a plausible solution to a potential

problem presented by sentences like (96):

(96) Zhangsan, [[*e* xihuan *e* de] ren] hen duo.
Zhangsan like DE man very many
a. 'Zhangsan, people who he likes are many.'
b. 'Zhangsan, people who like him are many.'

The sentence is ambiguous between a reading according to which the subject EC is bound by the topic *Zhangsan* and the object EC is relativized, and a reading according to which the binding configurations are exactly the opposite. In both cases, the binding relation between the topic and the EC violates Subjacency if the relation must be established by movement. It was suggested in section 3 that such apparent violations may be accounted for by assuming that the variable bound by the topic may start out as a pronominal and be converted into a variable. This is all right for the examples (78a) and (79a) considered above, as well as for the first reading of (96), in which the topic binds the subject EC. The problem arises with the second reading, where the topic appears to bind the object EC. The object variable cannot start out as a pronominal and be converted into a variable by the GCR, this possibility being jointly ruled out by DJR and the GCR. The problem can be avoided, however, if we assume that the object EC is first moved to a topic position within the relative clause, where it is then coindexed with *Zhangsan* in accordance with the GCR. Thus, given (96) at D-structure, Move α may apply and give rise to (97):

(97) Zhangsan, [[e_j e_i xihuan t_j de] ren] hen duo.
Zhangsan like DE man very many

In this structure the object EC is a *wh*-trace locally bound by e_j in topic position. The topic e_j may be coindexed with *Zhangsan* in accordance with the GCR, giving rise to the (b) reading of (96).

Another fact worth mentioning is that often adverbial clauses of time, place, etc., may seem to admit apparent pronominal ECs in object position. Kuroda (1965), for example, has observed that in Japanese, although an object gap in a complement clause may not refer to the matrix subject, an object gap in an adverbial clause may. Thus, in contrast to the (b) sentences of (25)–(26) considered earlier, sentences like the following are acceptable, each with an object gap that looks like a pronominal EC (see also Ohso (1976, 6)):

(98) John-wa$_i$ [Bill-ga e_i mita] hi ni wa byooki datta.
John Bill saw day sick was
'John$_i$ was sick on the day when Bill saw [him$_i$].'

(99) John-wa$_i$ [Bill-ga e_i sikatte] kara yoku benkyoosuru yooni natta.
John Bill scold since well study become
'John$_i$ has started to study harder since Bill scolded [him$_i$].'

Since the topic *John* is also the subject of the matrix clause in both (98) and (99), the

embedded object gap being coindexed with the topic also refers to the subject and looks like a pronominal EC. Similar observations may be made about Chinese:

(100) ni yi kan-wan e_i, jiu qing ba shu_i huan gei wo.
you once read-up then please BA book return to I
'As soon as you finish reading [it_i], please return the $book_i$ to me.'

(101) ruguo ni xie-chu yiben hao shu_i, wo yiding hui mai e_i.
if you write-out one good book I definitely will buy
'If you write a good $book_i$, I definitely will buy [it_i].'

In Portuguese, sentences like the following (Chao (1983b)) show the same fact:

(102) Se você não gosta desse $vestido_i$, não compra e_i.
if you don't like this dress don't buy
'If you don't like this $dress_i$, don't buy [it_i].'

In each of these sentences, an object EC in one clause is coindexed with an argument in the other. Since the sentences under consideration do not involve structures of complementation, we may again make use of the idea of a predication rule. In each case, one may assume that there is an empty operator adjoined to the minimal S node dominating the object EC (which need not be identified as a topic node), which binds the object variable and is coindexed with an external argument by predication.

Note that even without this assumption of an abstract operator, the fact that each of the object ECs in these sentences is coindexed with an argument does not itself establish the EC as a pronominal. This is because the coindexed argument does not c-command or bind the EC, as is clear at least with examples (100)–(102),[40] and this property of being coindexable with a non-c-commanding argument is consistent with the view that such an EC is a variable. As a variable, it is an R-expression in the sense of the binding theory, on a par with other R-expressions (lexical NPs and anaphoric epithets). It is well known that lexical NPs and anaphoric epithets have the two properties (a) that they may not be A-bound, and (b) that they may be coindexed with non-c-commanding arguments (Lasnik (1976)):

(103) a. *John likes the woman that met John.
b. The woman that met John likes John.

(104) a. *John said that the sissy would come.
b. Before I saw the sissy, John left.
c. Before I saw John, the sissy left.

I have claimed that an object EC is a variable and shown that it has the first property of lexical R-expressions in that it cannot be A-bound. The fact that it can be coindexed with a non-c-commanding argument should, in fact, come as no surprise, since it simply

[40] In the Japanese examples (98) and (99), if we assume that the topic *John* binds a subject EC in the matrix clause and that this EC occurs in the position immediately following the adverbial clause, then the object EC in the adverbial clause also is not c-commanded by the matrix subject EC.

testifies to what is naturally expected, that it also shares the second property of lexical R-expressions. If this reasoning is correct, then the occurrence of sentences like (100)–(102) may be taken as confirming evidence for the claim that an object EC is a variable.

The question remains, of course, what distinguishes discourse-oriented languages like Chinese, which allows sentences like (100)–(102), from non-discourse-oriented languages like English, which does not:

(105) *If you write a good book, I will definitely buy *e*.
(106) *If you don't like that dress, then don't buy *e*.

The relevant reason for the ill-formedness of (105) and (106) appears to be that they each contain a free variable, violating the condition that all variables must be Ā-bound (May (1977)).[41] To account for sentences like (100)–(102), the idea is then to say that languages like Chinese allow variables that may appear to be Ā-free. One way to execute this idea is simply to say that this condition of variable binding does not obtain in these languages. This simple execution may well be the right one, though it would have nothing to say about the German facts considered above. What I have done in this article is to suggest a more restrictive way of executing the idea. Instead of allowing free variables in potentially every position, I have suggested that all variables are bound, and that what can be free is only the empty operator that binds such a variable.

References

Aoun, J. (1982) *The Formal Nature of Anaphoric Relations,* Doctoral dissertation, MIT, Cambridge, Massachusetts.

Aoun, J. (forthcoming) *Generalized Binding,* Foris, Dordrecht.

Aoun, J. and D. Sportiche (1983) "On the Formal Theory of Government," *The Linguistic Review* 2, 211–236.

Borer, H. (1983) *Parametric Syntax,* Foris, Dordrecht.

Bresnan, J., ed. (1982) *The Mental Representation of Grammatical Relations,* MIT Press, Cambridge, Massachusetts.

Bresnan, J. and J. Grimshaw (1978) "The Syntax of Free Relatives in English," *Linguistic Inquiry* 9, 331–391.

Chao, W. (1981) "Pro Drop Languages and Non-obligatory Control," *University of Massachusetts Occasional Papers in Linguistics* 7, 46–74.

Chao, W. (1983a) "The Interpretation of Null Subjects: Brazilian Portuguese," *Cahiers Linguistiques d'Ottawa* 11, 59–74.

Chao, W. (1983b) "Discourse-based Null Elements and the Typology of Empty Categories," ms., University of Ottawa.

Chao, Y. R. (1968) *A Grammar of Spoken Chinese,* University of California Press, Berkeley, California.

Chomsky, N. (1973) "Conditions on Transformations," in S. Anderson and P. Kiparsky, eds., *A Festschrift for Morris Halle,* Holt, Rinehart and Winston, New York.

Chomsky, N. (1976) "Conditions on Rules of Grammar," *Linguistic Analysis* 2, 303–351.

[41] It seems that lexical R-expressions may also be Ā-bound, which is another property that they share with variables, although binding is required only of the latter:

(i) John, I like the sissy.
(ii) ?John, I like John.

Chomsky, N. (1977) "On *Wh*-Movement," in P. Culicover, T. Wasow, and A. Akmajian, eds., *Formal Syntax,* Academic Press, New York.

Chomsky, N. (1980) "On Binding," *Linguistic Inquiry* 11, 1–46.

Chomsky, N. (1981) *Lectures on Government and Binding,* Foris, Dordrecht.

Chomsky, N. (1982) *Some Concepts and Consequences of the Theory of Government and Binding,* MIT Press, Cambridge, Massachusetts.

Chomsky, N. and H. Lasnik (1977) "Filters and Control," *Linguistic Inquiry* 8, 425–504.

Cole, P. (1982) "On Null Pronominals," ms., University of Illinois, Urbana, Illinois.

Comrie, B. (1983) "Coreference in Discourse and Grammar," seminar lecture, USC, Los Angeles, California.

Evans, G. (1980) "Pronouns," *Linguistic Inquiry* 11, 337–362.

Gazdar, G. (1981) "Unbounded Dependencies and Coordinate Structure," *Linguistic Inquiry* 12, 155–184.

George, L. and J. Kornfilt (1981) "Finiteness and Boundedness in Turkish," in F. Heny, ed., *Binding and Filtering,* MIT Press, Cambridge, Massachusetts.

Gundel, J. (1980) "Zero Anaphora in Russian: A Case of Topic Prominence," *Proceedings of the Parasession on Pronouns and Anaphora,* Chicago Linguistic Society, University of Chicago, Chicago, Illinois, 139–146.

Hale, K. (1983) "Warlpiri and the Grammar of Non-configurational Languages," *Natural Language and Linguistic Theory* 1, 5–48.

Higginbotham, J. (1980) "Pronouns and Bound Variables," *Linguistic Inquiry* 11, 679–709.

Hornstein, N. and A. Weinberg (1981) "Case Theory and Preposition Stranding," *Linguistic Inquiry* 12, 55–91.

Huang, C.-T. J. (1982) *Logical Relations in Chinese and the Theory of Grammar,* Doctoral dissertation, MIT, Cambridge, Massachusetts.

Huang, C.-T. J. (1983) "A Note on Binding Theory," *Linguistic Inquiry* 14, 554–561.

Jaeggli, O. (1982) *Topics in Romance Syntax,* Foris, Dordrecht.

Jelinek, E. (1983) "Case, Government, and Configurationality," ms., University of Arizona, Tucson.

Kayne, R. (1981) "ECP Extensions," *Linguistic Inquiry* 12, 93–133.

Kayne, R. (1983) "Connectedness," *Linguistic Inquiry* 14, 223–249.

Kitagawa, C. and C. Ross (1982) "Prenominal Modification in Chinese and Japanese," *Linguistic Analysis* 9, 19–53.

Kuno, S. (1980) "Discourse Deletion," in S. Kuno, ed., *Harvard Studies in Syntax and Semantics* 3, 1–144.

Kuroda, Y. (1965) *Generative Grammatical Studies in the Japanese Language,* Doctoral dissertation, MIT, Cambridge, Massachusetts.

Lasnik, H. (1976) "Remarks on Coreference," *Linguistic Analysis* 2, 1–22.

Li, C. and S. Thompson (1976a) "Subject and Topic: A New Typology," in C. Li, ed., *Subject and Topic,* Academic Press, New York.

Li, C. and S. Thompson (1976b) "On the Issue of Word Order in a Synchronic Grammar: A Case against 'Movement Transformations'," *Lingua* 39, 169–181.

Li, C. and S. Thompson (1981) *Mandarin Chinese: A Functional Reference Grammar,* University of California Press, Berkeley, California.

Manzini, M. R. (1983) "On Control and Control Theory," *Linguistic Inquiry* 14, 421–446.

May, R. (1977) *The Grammar of Quantification,* Doctoral dissertation, MIT, Cambridge, Massachusetts.

McCloskey, J. and K. Hale (1982) "The Syntax of Inflection in Modern Irish," *Proceedings of the 13th Conference of the North Eastern Linguistic Society,* University of Massachusetts, Amherst, Massachusetts, 173–190.

McLuhan, M. (1964) *Understanding Media: The Extensions of Man,* McGraw-Hill, New York.
Mei, K. (1978) "Ba Zi Ju (On *Ba*-sentences in Modern Chinese)," *Wen Shi Zhe Xue Bao*, National Taiwan University, 27, 145–180.
Mohanan, K. P. (1983) "Functional and Anaphoric Control," *Linguistic Inquiry* 14, 641–674.
Nishigauchi, T. (1983) "Japanese LF: Subjacency vs. ECP," paper presented at the Seoul Workshop on Formal Grammar, Ewha Woman's University, Seoul.
Ohso, M. (1976) *A Study of Zero Pronominalization in Japanese,* Doctoral dissertation, Ohio State University, Columbus.
Perlmutter, D. (1971) *Deep and Surface Constraints in Syntax,* Holt, Rinehart and Winston, New York.
Pesetsky, D. (1982) *Paths and Categories,* Doctoral dissertation, MIT, Cambridge, Massachusetts.
Postal, P. (1971) *Crossover Phenomena,* Holt, Rinehart and Winston, New York.
Reinhart, T. (1976) *The Syntactic Domain of Anaphora,* Doctoral dissertation, MIT, Cambridge, Massachusetts.
Reinhart, T. (1982) "Pragmatics and Linguistics: An Analysis of Sentence Topics," Indiana University Linguistics Club, Bloomington.
Rizzi, L. (1982) *Issues in Italian Syntax,* Foris, Dordrecht.
Rosenbaum, P. S. (1967) *The Grammar of English Predicate Complement Constructions,* MIT Press, Cambridge, Massachusetts.
Ross, J. R. (1967) *Constraints on Variables in Syntax*, Doctoral dissertation, MIT, Cambridge, Massachusetts.
Ross, J. R. (1982) "Pronoun Deleting Processes in German," paper presented at the annual meeting of the Linguistic Society of America, San Diego, California.
Saito, M. (1982) "Comments on the Papers on Generative Syntax," in Y. Otsu et al., eds., *Studies in Generative Grammar and Language Acquisition,* International Christian University Press, Tokyo.
Suñer, M. (1983) "Pro_{arb}," *Linguistic Inquiry* 14, 188–191.
Taraldsen, T. (1978) "On the NIC, Vacuous Application, and the That-trace Filter," Indiana University Linguistics Club, Bloomington.
Teng, S. H. (1979) "Remarks on Cleft Sentences in Chinese," *Journal of Chinese Linguistics* 7, 101–114.
Thompson, S. (1973) "Transitivity and Some Problems with the *Ba*-construction in Mandarin Chinese," *Journal of Chinese Linguistics* 1, 208–221.
Tsao, F. (1977) *A Functional Study of Topic in Chinese: The First Step toward Discourse Analysis,* Doctoral dissertation, USC, Los Angeles, California.
Wasow, T. (1972) *Anaphoric Relations in English,* Doctoral dissertation, MIT, Cambridge, Massachusetts.
Williams, E. (1980) "Predication," *Linguistic Inquiry* 11, 203–238.
Yang, D. W. (1983) "The Extended Binding Theory for Anaphors," paper presented at the GLOW conference, University of York.
Zubizarreta, M.-L. (1983) "Theoretical Implications of Subject Extraction in Portuguese," *The Linguistic Review* 2, 79–96.

Department of Chinese Literature and Linguistics
and
Department of Foreign Languages
National Tsing Hua University
Hsinchu 300
Taiwan, ROC