

CHAPTER 1.

INTRODUCTION

Sometimes what is not said may convey as much meaning as what is. For example, if a woman says she has one brother, we quite readily infer that she has *only* one brother, though she has not actually said this. When a friend cries “Look out!” we conclude that the imperative is, or at least may be, directed at ourselves, though the friend has not specified who should follow his command. When we then turn and see a large truck bearing down on us, we further conclude that this truck is what we are supposed to look out *for* though this has been left similarly unspoken. In the extreme case, and using the most general definition of meaning, even refusing to answer a question conveys some information.

All of these phenomena are interesting and important parts of human communication, but some rely on more than just linguistic information. For a linguist attempting to model a grammar or develop a theory of grammar, one of the more interesting types of unspoken or *non-overt* information is found when some argument of a sentence or phrase is apparently missing. The grammar must account for these null arguments at some level and explain the difference between when they do and do not occur overtly.

This thesis describes certain ‘missing’ arguments in English, explaining how they fit into English grammar and how they relate to general theories of null arguments, and showing what type of meaning they convey in a discourse context. This information is incorporated into a formal discourse model, with the intention of making these discourse effects of null arguments in English accessible to more applied computational models.

1.1. WHAT NULL ARGUMENTS ARE

Natural languages have a variety of means for marking the contextual status of an entity. In particular, speakers can formally distinguish various levels of “givenness” or saliency that an entity may hold at a point in a discourse (cf. Prince 1981). A hearer uses these distinctions primarily to place information contained in an utterance into the appropriate discourse context.

In English, indefinite and definite determiners, personal pronouns, and reflexive anaphors are some of the tools used by speakers to mark different saliency levels. In examples (1) - (3) below the three sentences may all describe the same event; only the marked saliency of the entities represented by the subject and the object is changed.

- (1). A pilot flew a plane through the storm.
- (2). The pilot flew the plane through the storm.
- (3). He flew it through the storm.

Loosely speaking, the use of indefinite articles on pilot and plane in (1) may indicate that these entities are new to the discourse and perhaps also to the hearer. In (2), the definite article is a linguistic cue that the entities under discussion are identifiable, and not new to the hearer, either having been explicitly introduced into the discourse in some earlier utterance or being inferable in some way from the context. Finally, replacing the noun phrases referring to the pilot and the plane with pronouns in (3) indicates that these entities are not only previously identified, but also currently salient.¹

Although it is hard to imagine a reasonable natural language without tools of this sort, the exact nature of the marked saliency distinctions and the means to represent them may differ to some extent from language to language. Not all languages, for instance, have indefinite or definite determiners.

¹By ‘salient’ I mean that an entity has not just been previously mentioned but is also currently assumed to be in the hearer’s consciousness (Cf. Prince 1981).

Japanese is one example of a language with neither type. Optional adverbial number-marking, as shown in (4), is the closest functional approximation to indefinites.

- (4). Ringo o hitotu tabeta.
apple-obj one ate
“I ate an apple (lit.) one-ly” (Kuno 1973:26)

The lack of (in)definite determiners in a language does not entail that speakers of this language are incapable of conveying (in)definiteness in the semantic sense.² It simply means that, for whatever distinctions they do convey, they use other linguistic devices. In Japanese, for instance, the use of the morphological marker ‘-wa’, in addition to topic-marking the NP, indicates indirectly that the entity represented by that expression is not brand-new. Alternatively, Japanese speakers can also use demonstratives for this purpose (cf. Kuno 1973), as shown in (5).

- (5). Ano hito wa saikin doosite imasu ka?
that person TOP recently how is Q
“How is he doing these days? (Kuno 1973:282)

The absence of some methods for marking contextual status in a particular language may not be arbitrary. The structure of the language may not meet the necessary conditions for the existence of some linguistic device, or may heavily constrain the syntactic environments in which some device can occur.

For example, many languages, including English, do not allow long-distance anaphors, as shown in (6).

- (6). * Rob_i believes that Diana will be watching himself_i.

The binding conditions on anaphors in English force this exclusion, even though the construction could otherwise serve to disambiguate contexts such as the one in (7) below.

²In order to avoid the confusion that can arise from the ambiguous use of these terms, I will henceforth use (in)definite only to describe forms of expressions. ‘Brand-new’, ‘hearer-old/discourse-new’, and ‘discourse-old’, as defined in Prince 1981, will be used to describe the contextual status of an expression's referent.

- (7). a. Does John_j think Diana will see the surprise party guests arrive?
- b. He_j doesn't know, but ...
- Rob_i believes that Diana will be looking at him_{i/j} /*himself_i.

The point of this discussion is that null arguments, such as pro-drop and null objects, constitute another type of referential expression which may have a specifiable pragmatic/discourse function, and which has been argued to occur only in languages meeting certain strict structural conditions. Whether and to what extent they occur in English is still something of an open issue.

A clarification should be made here. The term *Null argument* is frequently used without a rigid definition, referring to one phenomenon here and another there. For my purposes, I would like to make use of this range of meanings in the following way -- henceforth a *null argument* is any argument in an utterance which does not have an overt phonological (or written) representation. In other words, I am starting with the broadest possible class of 'missing' subjects and objects, though various subtypes of null arguments may differ in form or function.³

1.1.1. Null Arguments In English

Relatively little attention has been given to English null arguments. The reason for their exclusion is primarily that they are not members of the *types* of null arguments most commonly considered.

The popular view is that some languages freely allow null arguments, while other languages fall into a small number of categories with limited null arguments, and still others allow none. Japanese,

³Note, however, that this definition does not include theory-specific elements like *traces* where an argument that is not overtly represented in its canonical sentence position but is overtly represented elsewhere in the sentence is argued to leave a phonologically null marker at another node (or nodes) in the sentence structure. Although the relationship between null arguments and these types of elements is an interesting study, it is orthogonal to this work (cf. Chung 1989).

Warlpiri and Russian, for example, are considered to be very free with their null arguments. Italian, Spanish, and even German are supposedly somewhere in the middle, while English and French are placed at the other extreme with, supposedly, no null arguments at all.

Under this view, there are implicit limitations on what constitutes a null argument. Specifically, an argument must be able to be non-overt in formal (even written) registers of the language under consideration, and the non-overtness of the argument must not be lexically dependent.

Most null arguments in English fail to meet one or the other of these implicit conditions and have therefore been excluded.⁴ For instance, example (8), which has a missing subject (i.e., *it*), is acceptable in conversational English but not in most written contexts, and example (9) demonstrates how two otherwise similar verbs in English may contrast with respect to the grammaticality of a null object. (A more complete overview of the English data is given in section 1.3.)

(8). Works every time.

(9). a. Henry finally ate 0/his dinner at midnight.

 b. Henry finally consumed *0/his dinner at midnight.

Despite these constraints, English null arguments have a lot in common with their better-known cousins and I will attempt to point out the similarities as well as the differences. As will be shown, the similarities do not lie entirely in the realm of sentential syntax. Null arguments across language types may be similar in a sense that is not apparent until we look beyond the level of the sentence. In order to argue for this similarity I must first demonstrate that it is possible to talk about language in a non-written context, and about intersentential structure.

⁴Imperatives are an exception and they will be considered separately.

1.1.2.

Null Arguments and Language Beyond Sentence-level

The definition of language that we linguists assume crucially affects what we study. These boundaries that we place on the domain of language are useful. They provide the necessary exactness needed for grammar-building. They may also, however, cause us to overlook some interesting phenomena.

For the last 30 years or more, the common assumption has been that language is divided (at least for the purposes of study and modeling) into a set of independent modules, most importantly syntax, semantics, and phonology.⁵ This assumption has led to a great deal of productive research and is still an important component of many theories of grammar. There is, however, another assumption that has been strongly tied to the assumption of modularity -- namely that these modules operate on the level of the sentence and only on that level. A model of language competence was said to be complete if it could accept all and only the sentences of a language. The facts of language in context were relegated to the nebulous realm of "language use" or "production", along with memory limitations, mental errors, rhetorical practices, and so forth.

Perhaps the motivation for this assumption was that it provided a relatively clear-cut area of study; the dividing lines between competence and use, and between language and general knowledge are more problematic if we do not accept the sentence boundary. Nonetheless, it is not obvious that sentences are the best units to use, even for the purpose of establishing boundaries.

In fact, the sentence is arguably an artifact of written language rather than a universal primitive. Certainly, a native speaker does not always speak in complete sentences, nor does she instinctively know that what is acceptable in spoken language is not necessarily acceptable in written language. It is even possible to compose sentences that obey writing constraints but do not obey the constraints of spoken

⁵Whether or not morphology is a separate module has been subject to debate.

language, as in (10) below. Therefore, despite its usefulness at one level, understanding the concept of 'sentence' in a language is not clearly equatable to having native linguistic competence in that language.

(10). The managing of an office by Peter is liked by John. (from Chafe 1992:18)

The tendency to define language based on sentences has led to a parallel tendency to look at what is acceptable in written language, the domain in which sentential units are most clear. However, since spoken language has been traditionally and is still currently thought of as the core of linguistic ability (cf. Aronoff (1994) for a recent discussion)), while written language is derivative, it is ultimately inappropriate to exclude constructions which occur primarily in conversational language. When there are differences between written and spoken language these must eventually be accounted for in the grammar.⁶ This task, broadly defined, may be somewhat more difficult than the already impressive goal of describing all that occurs in the domain of written language, yet we can look at individual conversational phenomena such as null subjects in English with relative ease.

There is precedence for this approach. Most research has at this point gone beyond the basic 'language-as-sentence' definition. As will be discussed in Chapter 2, it now seems clear that grammars that hope to represent conversational data must accept minimal units smaller than sentences. More saliently, certain constructions that had been rejected as ungrammatical since they tend not to occur in formal language, have been re-assessed as grammatical components of English. For example, Kroch (1981) has demonstrated, relying largely on data from spoken language, that resumptive pronouns play a legitimate role in English -- in both spoken and written language, and Prince (1984) has shown that left-dislocation is a common and pragmatically well-defined construction in spoken English.

Once we move beyond the idea that a language is a set of sentences, we are also able to consider the possibility that our linguistic competence includes an awareness of units larger than a single sentence.

⁶The exception to this is when non-linguistic cognitive factors such as memory constraints can account for the difference.

Prince (1988) describes some criteria for including discourse phenomena in the study of language. I discuss and expand this idea further in Chapter 5 of this thesis. But again, there is already theoretical work that relies on this perspective. As an example, research by Heim (1982) and Kamp (1981) on the semantic scope of indefinites relies on the possibility of a domain larger than a sentence.

The potential benefits of considering the larger class of null arguments across languages are clear. A non-exhaustive list includes: helping us to clarify the notion of ‘zero pronoun’ in so-called ‘discourse-oriented’ languages (described in Section 1.2.2 below), providing a broader picture for distinguishing (or collapsing) the theories of *pro* and PRO, bringing other apparently idiosyncratic pro-drop languages (such as Yiddish) into a unified account of pro-drop, and showing how some apparent pragmatic inferences may be represented in grammatical structure.⁷

1.1.3. Issues

The broad definition of null arguments stated earlier leaves open four major issues. First, what methods can be used to determine the presence of an argument in the absence of an overt representation; second, at what level(s) of the grammar are the various types of null arguments represented; third, what, if any, particular grammatical conditions do null arguments require; and fourth, what special pragmatic or discourse functions do null arguments have.

These issues have been addressed to one degree or another for various types of null arguments in languages other than English. In the remainder of this introductory chapter I discuss work that has been done on null arguments cross-linguistically, addressing the extent to which generalizations have been and can be made, give a general overview of the types of null arguments found in English, and describe how this thesis is organized.

⁷ Some of these ideas will be addressed further later in this thesis.

1.2. A REVIEW OF CROSS-LINGUISTIC RESEARCH ON NULL ARGUMENTS

In this section, I discuss the various types of null arguments that have been studied in languages other than English, detailing the theories that have been proposed to account for these phenomena. Each of these theories has both strengths and weaknesses. I attempt to point these out, and I suggest an alternative perspective on cross-linguistic generalization.

1.2.1. Pro-drop

Pro-drop is a term from the syntax literature used to refer to a missing subject which is treated as an pronominal empty category, *pro*, in the subject position of a tensed clause and which is licensed by particular syntactic or morphological properties of certain languages. It is also considered by some to be a parameter of universal grammar, patterning with subject-inversion and lack of that-trace effects. Some more or less clear examples of pro-drop languages include Spanish, Italian, Catalan, and Hebrew. The null subjects in these languages are claimed to be of the same type, all licensed by syntax and all somehow marking the person and number of the missing subject. For example, in the Spanish sentence in (11), the missing subject in each case is recoverable from the form of the verb.

- (11). Compran/compramos unos recuerdos.
buy-PRES-3plur/buy-PRES-1plur some souvenirs
“(They/we) are buying souvenirs.”

The use of pro-drop in these languages is most often described as equivalent to the use of unstressed pronouns in languages like English. In other words, pro-drop is the most common type of ‘pronominal’ subject in these languages (cf. Vallduví 1987, Di Eugenio 1990).

There have been several proposals about the exact nature of the syntactic licensing of pro-drop. The first of these is that, though subject case must always be assigned, in languages with ‘strong’ inflection, the subject case may be absorbed by the inflective morphology, eliminating the need for an overt subject. In languages with weak inflection, the case can only be assigned to an overt subject. One version of this proposal also accounts for certain ‘discourse-oriented’ languages (discussed below) which do not have inflection by suggesting that they do not have IPs (cf. Speas 1994).

Another proposal is that in certain languages the strong inflection is, in fact the subject, and an overt ‘subject’ is actually in an adjunct position (cf. Jelinek 1984). This argument has been used to explain examples like (12) in Spanish, where the inflectional features on the overt subject do not match those on the verb.

- (12). Las mujeres tenemos esperanza. (Jelinek 1984:48)
the women have-PRES-1plur hope
“We women have hope.”

It has also been suggested that similar constraints may allow *pro* to occur in the object position of some languages. However, null objects are not freely found in all pro-drop languages, and the idea of *pro* occurring in object position has not been widely accepted. An alternative theory is that there is a pronominal empty category in TOPIC position in these cases, while the object position has a variable (again null) indexed with this topic (cf. Huang 1984, Authier 1989, Speas 1990). This theory predicts different behavior than does a pronominal account of null arguments. Empirical evidence has been presented against the variable analysis as well as for it, leading some linguists to maintain support for *pro* objects, at least in some pro-drop languages (cf. Farrell 1984).

that there is a “Zero Topic Assignment” effect in which, under the right conditions, a null argument may act as if it is topic-marked.

Discourse-oriented languages are also noted to use anaphors that are not syntactically bound. An example in Korean is given below:

- (14). a. John-i salam-iŋ ponae-ess-ni?
 -NOM -ACC send-PAST-Q
 “Did John send the man?”
- b. ani, caki-ka cikcep o-ess-ta.
 no self-NOM in-person come-PAST-DEC
 “No, self came in person.” (Huang 1984:57)

The other property that has been informally associated with discourse-oriented languages is the “rich semantic specification of arguments.” (Kameyama 1985:32). Kameyama lists four particular types of semantic specification that seem more prominent in Japanese and Chinese. Perhaps the most definitive of these is what she refers to as the “strict animacy constraint.”⁹ This constraint in colloquial Japanese prevents non-animate entities from being the subject of transitive constructions. The other types Kameyama cites include a more distinct set of transitive verbs for [+human] vs. [-human] objects in Chinese, subtler distinctions in semantic fields, particularly in verb meanings or contexts, and honorificity features on verbs. Kameyama notes however that the extent to which this property extends to discourse-oriented languages other than Japanese and Chinese is yet to be seen.

The null arguments in discourse-oriented languages have been argued to be phonologically null (zero) pronominals that do not require overt person-number marking and whose antecedents are determined instead by discourse features. In other words, the argument is that zero pronominals are simply permissible in discourse-oriented languages without any special requirements, and that zero

⁹This property is described for Japanese in Kuno (1973).

availability of null anaphora for all arguments, these non-configurational characteristics most commonly include free word order and the use of syntactically discontinuous expressions. (cf. Hale 1983).¹¹

A relatively well-known example of this type of language is the Australian aboriginal language, Warlpiri. Warlpiri makes extensive use of null anaphora. As in (15), the features of the null arguments are generally marked on the auxiliary, with third person singular being the unmarked case.

- (15). Punta- rni kapi-rna-ngku.
 divest NONPAST FUT-1subj- 2obj
 “(I) will take (him/her/it) away from (you).” (Hale 1983:19)

Although null anaphora often co-occurs with person-number marking in non-configurational languages, Hale points out that this is not always the case. For instance, unlike Warlpiri, the Arandic languages of Central Australia lack a person-number system yet they too allow null anaphora. This is illustrated in the following example from Kaititj:

- (16). Weye wampere atye are-nhe, eylpwere-rle
 meat possum I(ERG) see-PAST hollow LOC
 anteyanenge- warle, elepe- le- lke atye arte-nherre,
 sitting COMP axe INST AFTER I(ERG) chop PAST,
 eylpwere, alarre- nhe atye.
 hollow kill PAST I(ERG)
 “I saw a possum sitting in a hollow tree: so then I chopped it, the hollow tree
 and killed (it).” (Hale 1983:17:fn9)

Hale (1983) suggests that there is a configurationality parameter which allows for variation in the domain of the projection principle. Roughly speaking, non-configurational languages only require the

¹¹Not all non-configurational languages have every one of these characteristics. It should be noted in particular that even null anaphora is not universal. (cf. Hale 1983)

projection principle to apply at LS. Certain superficial characteristics, such as those mentioned above, become possible but not necessary under these conditions. As applied to null anaphora, this can be interpreted as meaning that argument positions do not have to project to surface structure. Hale suggests that null anaphora in Japanese and other discourse-oriented languages could also potentially be explained in this way. In fact, contrasting with the previous discussion, Japanese has at times been argued to be a non-configurational language. (cf. Chomsky 1981) This analysis raises the issue of the reality of the contrast between discourse-oriented and non-configurational languages.

An alternative proposal in Simpson (1983), is that non-configurational languages, or at least Warlpiri-like languages, allow null pronominals mainly because the AUX features meet a “completeness constraint”, i.e., the features of the null argument are recoverable from these AUX features. Though this proposal allows for a stronger similarity between null anaphora in non-configurational languages and pro-drop, the parallel is not perfect. Simpson notes that the case is somewhat different in nonfinite clauses in Warlpiri. “...even in non-finite clauses, it is possible to have null pronominals with non-third person referents, providing this information is recoverable from the context.” Simpson's proposal also leaves open the issue of those non-configurational languages which have null anaphora but do not have person-number marking.

1.2.4. Arbitrary and indefinite null objects

Some languages allow only a restricted use of null objects, occurring in sentences with ‘generic’ time reference provided they are also given an arbitrary interpretation. Italian and French, though they differ with respect to subject pro-drop, are both argued to have these arbitrary null objects (cf. Rizzi 1986, Bouchard 1987, Authier 1989a). (17) demonstrates the use of an arbitrary null object in Italian.

(17). Questo conduce a concludere quanto segue. (Rizzi 1986)

This leads to conclude what follows

“This leads (one) to conclude what follows.”

Rizzi (1986) and Authier (1989a) both argue that these null objects are structurally present in the sentences in which they occur, though they differ in some significant details. Rizzi treats them essentially as null pronouns with the feature [+arb], while Authier argues that they are variables bound at LF by overt or non-overt adverbs of quantification adjoined to the clause whose tense selects them. In contrast to both of these theories, however, Bouchard (1987) allows for the possibility that these null objects do not project out of the lexicon.¹²

Other specialized or idiosyncratic types of null objects have also been found. For example, Campos (1986) shows that indefinite direct objects may be dropped in Spanish, as in (18) below.

- (18). a. ¿Compraste café?
buy-past-2sing coffee
“Did you buy coffee?”
- b. Si, compré.
yes buy-past-1sing
“Yes, I bought (some).”

Campos argues that this is a variable position (a trace of OP). In this case, he is motivated to treat the null argument as something other than a null pronoun because it is indefinite, which would conflict with the Chomsky (1982) proposal that *pro* must have definite reference. Campos also argues that the syntactic behavior of these null objects is consistent with that of a variable.

¹²Bouchard also argues that the same phenomenon occurs in English, which I discuss in Chapter 3.

1.2.5. PRO

Of course, another ‘special’ null argument that has received significant attention is the missing subject in non-finite clauses. This is a phenomenon common to many languages, including English, as shown in the parallel examples in (19) below.

- (19).
- a. John tried to help me.
 - b. Jean a essayé de m'aider. (fr.)
tried to 1sg-help
 - c. Juan trató de ayudarme (Sp.)
tried to help-1sg
 - d. zhangsan shefa bangmang wo (Chin. - Huang 1989:197)
try help I

One theory is that these missing subjects are a unique type of empty category, PRO, with features that differ from other empty category types such as *pro*, traces and variables. The normal binding conditions on pronouns or anaphors do not consistently apply to this empty category, instead a separate theory of *control* attempts to explain how PRO is assigned an interpretation.

In an alternative approach, within the lexical-functional grammar framework, non-finite clauses have been treated as lacking an NP position for a subject, though the subject is included in a functional representation of the sentence.¹³ In this theory, *functional control* is lexically-specified in the entry for the matrix verb, linking the subject of the non-finite clause to an argument in the matrix clause (cf. Bresnan 1982).

¹³More precisely, the subject in these constructions is represented in ‘f-structure’ but not in ‘c-structure’.

These theories both maintain a clear distinction between PRO and other null subjects. There have, however, been proposed modifications to these theories which blur or eliminate this distinction. For instance, Huang (1989) proposes a generalized control theory intended to account for all null subjects in Chinese. Specifically, he suggests that *all* null pronominals get controlled within their control domains *if they have one*. Similarly, Bouchard (1985) argues that PRO, instead of bearing both [+pron] and [+anaph] features is a null pronoun in some environments and a null anaphor in others; In other words, it is not a special type of empty category, it simply occurs in two different types of environments.

1.2.6. Some (even) more idiosyncratic cases

There are a number of languages that exhibit null arguments that do not fit into any of the categories described above. This section should be considered just a small sampling of these languages.

In German for instance, only expletives may be null, as shown in (20).¹⁴ More precisely, under certain conditions, expletives *must* be null, as shown in (21) (cf. Safir 1985).

- (20). a. *I_t.
eat-3sing
 “(He/She) eats.”
- b. Er sagte, daß getanzt wurde.
he said that danced was
 “He said that there was dancing.” (Safir 1985:197)
- (21). Gestern wurde *(es) erklärt, daß ...
yesterday was it explained that
 “It was explained yesterday that ...” (Safir 1985:216)

¹⁴Even the German equivalent of the English weather it apparently must be overt.

- c. Badarft bentshn goyml.
 (She) works too hard.” (Prince 1994:2)
- (23). Take, meynst mist an emes?
 “Indeed, (you) really mean it?” (Prince 1994:4)
- (24). O, ven ikh volt es gekent makhn azoy, az aykh zol zayn gut bay undz.
 “Oh, if I could have made it so that (it) was pleasant for you here.”
 (Prince 1994:7)

Prince (1994) demonstrates that null arguments are not a unitary phenomenon in Yiddish. In addition to the different syntactic constraints described above, her data analysis indicates a diverse set of phenomena with different discourse constraints. In particular, in her examination of null subjects that represent ‘nominally-introduced’ discourse entities (i.e., those null subjects whose referents have already been overtly introduced into the discourse context), the high frequency of missing second person singular subjects (Prince's “Du-Drop”), is attributable to its relative insensitivity to other discourse constraints. All other null subjects in this class (i.e. 1sing/plur, 2plur, and 3sing/plur) are much rarer, being likely to occur only when their referent is the entity which the discourse just before that point was most centrally about.¹⁵ Finally, Prince notes that Yiddish also has null objects. Unlike the null subjects, they are not restricted to colloquial language and are not licensed by overt agreement. Unfortunately, to the best of my knowledge no work has been done to determine exactly how extensive or how pronoun-like these null objects are.

Null pronouns also occur in a variety of positions in Chamorro. Some of these positions involve rich agreement and some do not. According to Chung (1984), there is no person/number morphology to identify null objects, and yet they occur. These null objects do, in fact, behave like pronouns, not like

¹⁵This entity may be loosely thought of as the discourse topic, though Prince's study was actually conducted more formally within the framework of centering theory, a theory of local discourse structure which is discussed in some detail later in this thesis. In centering terms, these null subjects only occur in ‘Continue’ transitions.

variables. An interesting Chamorro constraint is that, if only one NP is overt, the overt argument must be the object.

Like Chamorro, Brazilian Portuguese (BP) allows null objects though there are no morphological features to make them syntactically ‘recoverable’. Farrell (1990) shows that, as with Chamorro, these null objects are pronominal in nature, but, unlike Chamorro, only third person objects may be null. In addition, there is apparently a marked preference for null inanimate objects and overt animate ones. Farrell suggests that there is an appropriateness constraint on null objects based on animacy, with inanimate objects being most appropriate and first/second person objects being completely inappropriate.

Fijian is another interesting example because it marks the ‘suppression’ of an obligatory argument with reduplication. Jonas and Lathroum (1992) note this pattern and offer the example below.

- (25).
- a. e vuli-ca [a fika]_o [o Maria]_s
“Mary is studying math.”
 - b. e vuli.vuli. [o Maria]_s
“Mary is studying and studying.”

Only truly transitive verbs reduplicate, not intransitives or unaccusatives. Fijian has the additional intriguing property, again described in Jonas and Lathroum (1992), that not *all* obligatory arguments of a verb may be suppressed. So, when reduplication is applied to what they call a “Type E” verb (a verb with an unaccusative analysis), the missing object has definite reference, as in (26)b below.

- (26).
- a. e kama-ca [a isulu]_o [o Maria]_s
“Mary is burning the garment”
 - b. e kama.kama [o Maria]_s
“Mary is burning and burning (it).”
 - c. e kama [a isulu]_s
“The garment is burning.”

Finally, Lillo-Martin (1986) argues that American Sign Language (ASL) has two different types of null arguments, the type available being dependent on the verb. The first type, which occurs with verbs that can spatially mark agreement, is obligatory when this marking occurs, and behaves like a pronominal. In this sense, says Lillo-Martin, it is like null arguments in Irish, which are also obligatory when certain morphology is used.

The other type of null argument in ASL occurs with verbs that can not mark agreement (what Lillo-Martin calls “plain verbs”). Lillo-Martin argues that ASL may be considered a discourse-oriented language, and that Huang's theory of a null TOPIC-variable construction holds for null objects of these plain verbs, while the null subjects in these cases would presumably be treated like the null subjects of other discourse-oriented languages. It seems to me that this raises an interesting issue, i.e., if Chamorro is a discourse-oriented language, why does subject agreement get marked for *any* verb?

1.2.7. A unifying theory?

It should be clear from this discussion of existing theories that the cross-linguistic data is rich and diverse. None of the theories above even attempt to capture all the data that has been described. Of course, efforts have been made to form a typology of languages more completely or in other ways with respect to the availability of null arguments. One good example is Kameyama (1985).¹⁶ Considering two characteristics, the need for overt pronouns and the presence of bound morphemes with argument features, she argues for four basic types of null subjects: (I) languages which require free pronouns and have bound morphemes indicating argument features (eg., German, French, English), (II) languages which have obligatory bound morphology in the form of either inflections, affixes or clitics but do not require overt free pronouns as well (eg., Italian, Spanish, Warlpiri, Chichewa), (III) languages which require neither

¹⁶It should be noted though that Kameyama appears to be influenced by phylogenetic and geographic patterns as well as by parametric possibilities.

(Japanese, Chinese, Malayalam) and (IV) languages which require free pronominals and do not have bound morphology (eg., Afrikaans, Swedish).

Kameyama allows for a significant amount of variation within these types. For instance English is classified with languages which require both overt subjects and overt bound subject morphology because of its solitary 3rd person singular marking, contrasting it with Afrikaans, which has no subject-verb agreement, and equating it with German which has a richer inflectional system but nevertheless generally requires overt subjects. Kameyama also creates a special subtype of the type IV languages, to include a group of languages which allow null third person subjects and have no overt “dummy” subject (though it's not clear why these languages are not equally acceptable as members of a subclass of type III.)

This typology is designed to be enumerative, ie. every language of the world by definition is assigned to one of these types, but it is not an explanatory one. First, it looks only at the surface similarities, eg., it does not distinguish between languages that represent null arguments differently or allow them under different conditions. Secondly, it doesn't allow for degrees to which overt pronominals are required or the relative strength or weakness of the bound morphology used.

It seems beneficial to question whether any simple set of parameters is sufficient to explain the full universe of possibilities for null arguments. An alternative hypothesis is that there is, in fact, a continuum from null argument to non-null argument languages rather than a number of preset language types. Under this analysis, the grammatical properties of a particular language, as well as other choices that the language makes about representing saliency and other discourse/pragmatic features, may affect the likelihood of null arguments in that language but not predetermine it. The only absolute constraint on null arguments would then be related to the issue of *interpretability*, a necessary component of all aspects of a language. A language may, in principle, obey this constraint by using null arguments that have one (or more) of the characteristics below:¹⁷

¹⁷I am not making the strong claim here that this is an exhaustive list. My main point is that there are multiple ways to satisfy the interpretability condition. Note also that the interpretability

- (27). **Means for Achieving Interpretability of Null Pronouns**
- i. Recoverable from grammatical features in sentence
 - ii. Recoverable from discourse/situational context
 - iii. Recoverable because of limited feature variation in null pronouns
 - iv. Arbitrary in reference
 - v. Expletive

We can also ask whether we can identify certain general discourse effects associated with the types of null arguments available in a language and the way it satisfies interpretability, or if the function of null arguments is partially or completely independent of these patterns.¹⁸

1.3. INTRODUCTION TO THE ENGLISH DATA

The following are examples of the types of null arguments found in English, all of which will be discussed in this thesis.¹⁹

- (28). Should've left while I had the chance. (1st person)
- (29). Arrived yesterday. (1st, 2nd or 3rd person)
- (30). Certainly is an unusual painting style. (3rd person)
- (31). Looks like rain. (weather "it")
- (32). Seems to be working. (pleonastic "it")

condition is, roughly speaking, a broadened and more robust version of the recoverability condition, but I avoid this term both to avoid confusion and because it carries connotations of a syntactic/morphological process.

¹⁸See Prince (1996?) for an interesting, related discussion of the independence of (the instantiated form of) a grammatical 'construction' in a particular language and its discourse function.

¹⁹The reader may notice that I have left out the null arguments of nominalizations such as "the (boy's) destruction (of the sandcastle)", as well as the passive, tough and middle constructions (and probably other types as well). Although it would be interesting to see if the ideas expressed here extend to these cases, particularly in the discussion of discourse theory in Chapter 5, a treatment of each of these types involves more than just a null argument in its canonical position and is therefore beyond the scope of this current work. I leave this as an area of future study.

- (33). Drive on the right side of the street. (imperative subject)
- (34). Finish your project yet? (Aux - subj)
- (35). John thinks that feeding himself will be difficult. (pronominal PRO)
- (36). The children want to go. (anaphoric PRO)
- (37). John called. (object)
- (38). Everyone drank at Bill's party. (object)
- (39). The animals always know first. (event reference)
- (40). This insurance policy only protects against theft. (arbitrary null object)
- (41). Insert in slot B. (objects in instructional registers)

Superficially, these examples look very similar to the types of constructions found in the languages mentioned above. In particular, examples (28) - (32) look identical to examples of pro-drop in other languages. (33) and (34) are examples that also look similar, though not identical. (33) is, of course, an imperative construction with a null subject, distinguishable on the surface from pro-drop by verb form. (34) also looks like pro-drop but the auxiliary verb is missing as well. Examples (35)-(36) are instances of English PRO. Finally, examples (37)-(41) could a priori be hypothesized to be zero pronominal objects.

Of course, many of these examples may, under closer observation, reveal themselves to be instances of dramatically different grammatical forms, not at all *syntactically* like their cross-linguistic equivalents. Others may turn out to have very similar syntactic effects, and even those which rely on different grammatical explanations may play similar roles in discourse structure. The major goal for this thesis is to characterize both the grammatical and the discourse properties of the null arguments in these English examples and point out the relation between the English constructions and the null argument constructions found in other languages.

1.4.

OUTLINE OF THE THESIS

Chapter 2 deals with null subjects in English. I argue that, while deletion of phonological material in sentence-initial position does occur in conversational English there is evidence for syntactically-present null subjects as well. Using a large corpus of spoken English, I demonstrate the existence of unique discourse constraints on null subjects in English. In particular I show that null subjects affect discourse structure at a both a hierarchical and a non-hierarchical level.

Null objects are considered in Chapter 3. Treating the different types of English null objects individually, I show that they are lexically-present but not syntactically-present. I argue that information in lexical structure can be distinguished from world knowledge based on discourse effects and I discuss the discourse properties of the different types of null objects.

Chapter 4 addresses the issue of variation, considering differences between Standard American English and null arguments in other dialects and registers of English. The special characteristics of recipes and telegraphs are also examined.

In Chapter 5, I further motivate the inclusion of discourse structure as a part of the study of linguistic competence, outline the types of discourse theories that are available for this purpose, and incorporate null arguments in English into a theory of local discourse structure. In order to accomplish this last goal, I define a modified version of Centering Theory that relies on the building blocks of lexical structure.

Finally, Chapter 6 reviews some of the results of earlier chapters and compares my treatment of English null arguments with existing theories of null arguments, discussing differences and similarities, and offering some suggestions about a more unified approach to null arguments. It concludes by briefly describing a number of areas for future research.