CHAPTER 5. A REVISED CENTERING MODEL FOR DISCOURSE STRUCTURE

In previous chapters, I have discussed the general grammatical and contextual constraints on null arguments in English. I have also introduced the notion of discourse packaging to explain the contextual constraints on conversational null subjects. In this chapter, I describe why it is important to consider English null arguments in a theory of local discourse structure, and explain why centering theory provides a good environment for this purpose. I then modify an existing centering model to incorporate the facts of English null arguments, in particular showing how the lexical conceptual structures used to treat null objects in Chapter 3 can be used as input.

5.1. THE NEED FOR A LINGUISTIC THEORY OF DISCOURSE STRUCTURE

As mentioned above, this chapter is about incorporating English null arguments into a theory of local discourse coherence. Because treatments of the syntactic/lexical properties of linguistic phenomena are not often joined with an analysis of their discourse properties, I would like to motivate this chapter with a brief discussion of the importance of discourse theory in the study of formal linguistics.

As mentioned in Chapter 1, I am starting with the assumption that knowledge of a language is more than knowledge of individual sentences. In other words, I assume that discourse information is crucial to a complete theory of language. In doing so, I am following Prince (1988) who argues that the existence of arbitrary and language-specific syntactic and referential options for conveying a proposition requires a level of linguistic competence beyond sentential syntax and semantics. Sentential models of

linguistic competence are unequipped to explain the existence of and the difference between multiple sentence forms with the same semantic interpretation. For instance, English topicalizations are semantically indistinguishable from canonical sentences unless we consider the discourse constraints on when they may occur (see also Prince 1981b.) Similarly, Prince argues, sentential grammars alone are not capable of constraining the use of definite and indefinite NPs.

There are several additional reasons for assuming that linguistic competence must be modeled beyond the level of the sentence. First, sentential grammars rely on the artifactual boundaries of written language. In some respects, this is a (short-term) advantage. The boundaries may be too small but they nonetheless provide a well-defined range of linguistic phenomena for a model of language to explain. In fact, this approach has been taken by generative grammarians for years with a great deal of success. However, the long-term disadvantages are also obvious. When one starts with a particular definition of language, any phenomena that do no fit into that definition will generally be ignored. If that definition is too narrow, then crucial data may be lost. Also, choosing to define language in terms of sentences in particular automatically includes a bias towards the type of language that one has been trained to consider 'proper' as opposed to what one knows through the initial process of first language acquisition.[1] This argument alone simply motivates a grammar that is less restricted than a sentential grammar; it does not necessarily require a grammar that includes discourse principles. However, once we accept that a language is not a set of sentences, we are free to explore broader possibilities.

Second, additional phenomena have been found that require at least a limited extension of sentential grammars. For example, sentential grammars can not completely account for the determination of pronoun co-reference, the scope of quantifiers, or the use of discourse deixis (Webber 1990). In addition, as discussed in earlier chapters, English null arguments provide more evidence that knowledge

---

[1]This doesn't mean that certain obvious or famous differences won't be noticed and therefore compensated for. (eg., split infinitives, dangling participles, and ain't.) But there is no way to judge if there are any aspects of our notion of `sentence' that are so ingrained that we do not recognize them as secondary rules that are not a part of our natural, conversational, language. The biases that have existed against resumptive pronouns and left-dislocations are examples of this type of error (but cf. Kroch 1981.)

of a language consists of more than a grammar for producing and interpreting sentences. I have shown that null subjects play an active role in conversational English, though they have received little attention in the past due to their rarity in written or 'formal' English. I have also shown that the presence of implicit null objects in English may not be distinguishable from truly intransitive constructions without an examination of extra-sentential information.

At some point the boundaries of linguistic theory must be extended to attempt to incorporate at least the types of phenomena mentioned above, and perhaps other constraints considered in the field of linguistic pragmatics as well.[2] The only alternative would be to keep sentential grammars and make the claim that all other phenomena are outside of the proper study of formal linguistics. This approach would implicitly claim that these phenomena are somehow not 'linguistic' and can not be handled properly without a complete model of, say, knowledge representation and its relation to language. It would also implicitly claim that there is no knowable boundary between language and knowledge except for the boundary of 'sentences' (or their undefined oral equivalent).[3] It seems clear, however, that it is, at the very least, the role of linguistics to test these claims. Linguists must, ultimately, attempt to discover the full range of phenomena that can be handled in a linguistic theory.

I therefore assume that these phenomena motivate the assumption that models of language require a discourse module. The real issue then is what this module should look like.

---

[2]I think the qualities of arbitrariness and language-specificity, as described in Prince (1988), may be used as a good heuristic for distinguishing linguistic knowledge from non-linguistic knowledge. More specifically, I think that phenomena that do not have these properties most likely involve more general reasoning processes, but these properties may not be the only common features of linguistic phenomena. (For example, the proper form for greeting important persons may be both arbitrary and language-specific, but knowing the language and knowing the specific context is not sufficient; knowledge of cultural/societal norms is also required.)

[3]'Utterances' are not the oral equivalent since they are interpreted in context and are dependent on the existence of discourse theories such as those I will discuss.

5.2. TYPES OF DISCOURSE THEORIES

Fortunately, in recent years, formal and computational linguists have begun to recognize the need for theories that extend beyond sentences. Several models now exist that attempt to handle at least some of the linguistic phenomena mentioned in the previous section. These models arise from a variety of different motivations and include many different approaches to defining and representing discourse information.

5.2.1. Classifying Discourse Theories

A number of attempts have been made to classify the types of information that may be contained in these models. An early approach to classifying discourse information is found in Halliday and Hasan (1976). Halliday and Hasan argue that there is a difference between 'cohesion' and 'coherence' in discourse. They claim that, "Cohesion occurs where the INTERPRETATION of some element in the discourse is dependent on that of another."(1976:4) Their notion of cohesion includes a definition of what constitute cohesive elements in a discourse. Specifically, they argue that there just four types of cohesive relations: Reference (situational and textual), Substitution, Ellipsis, and Conjunction. In other words, a discourse is more or less cohesive depending on how much use it makes of these relations. 'Coherence', to Halliday and Hasan, is concerned with logical (i.e., nonstructural) connections between pieces of discourse information.

In more recent work, Johnson-Laird (1983) argues that there is a difference between 'coherence' and 'plausibility'. Coherence, to Johnson-Laird, is concerned with linguistic information about co-reference and consistency, where consistency refers to the problems that would arise in interpretating a discourse which ascribes conflicting properties to an entity. Plausibility, in his view, depends on world conditions; a plausible discourse is constrasted with 'structured nonsense'.

While there is some relationship between Johnson-Laird's notion of coherence and Halliday and Hasan's notion of cohesion, there are some crucial differences. First of all, Johnson-Laird, coming from the perspective of building mental models of discourse, does not limit coherence to a specifically required set of structural devices. In addition, it differs from Halliday and Hasan's notion of cohesion by the inclusion of the concept of consistency. (Halliday and Hasan would treat this as a coherence factor, along with at least part of what Johnson-Laird considers to be plausibility issues.)

At this point, it should be clear to the reader that there are not yet any universally accepted terms or methods for discussing discourse structure. While Halliday and Hasan are concerned with separating structural and non-structural aspects of discourse, Johnson-Laird seems more concerned with separating linguistic and non-linguistic aspects of discourse. I consider the latter distinction more fundamental for our current purposes although, as will be shown below, some discourse theories blur this distinction.[4]

Walker, Iida and Cote (1994) offer yet another useful classification method. A discourse model may look at the role of semantic information (including the broader semantics of the discourse domain as well as the semantics of individual sentences) or it may focus on syntactic information. (Though they argue that both types of information are crucial and mutually constraining for the interpretation of utterances in a discourse context.) I will also make use of this distinction, within the bounds of linguistic models.

Finally, Grosz and Sidner (1986) distinguish between 'global' discourse structure and 'local' discourse structure. Since a model of global discourse structure is concerned with attentional state changes and the evolving intentions of speakers and hearers over large pieces of discourse rather than with the relationships between individual utterances, it is not relevant to the null argument data considered in this thesis. Though Grosz and Sidner (1986) provides a suggestion for determining local

---

[4]Johnson-Laird suggests that work such as that in Minsky (1975) and Schank and Abelson (1977) on *scripts* is concerned primarily with discourse plausibility. I will henceforth accept this characterization and will give these models no further consideration here.

discourse segments, it does not concern itself with phenomena that occur within a single segment.[5] All the models considered below deal with local discourse structure (though perhaps not explicitly or exclusively.)

5.2.2. Some Existing Models of Local Discourse Structure

In formal semantics, work by Heim (1982) on 'File Change Semantics' and by Kamp (1981) on 'Discourse Representation Structures' has come to be jointly referred to as Discourse Representation Theory (DRT). I will begin with a discussion of their work, and because they are quite similar models, I will treat them together.

Both Heim and Kamp are concerned with the relationship between entities in adjacent utterances and can therefore both be considered as models of local discourse. Both are concerned primarily with semantic information. They examine how the semantic representation of an utterance can add new discourse entities (referred to as file cards and reference markers respectively) or limit the availability of certain discourse entities to serve as antecedents for subsequent co-reference. Specifically, they both look at how indefiniteness and definiteness in various semantic structures such as so-called "donkey sentences" affect the creation of and continuing availability of discourse entities (what Heim calls their 'lifespan'.) A distinguishing feature of these models is that they are concerned with the previously unexamined extra-sentential constraints on non-coreference rather than with the determination of pronominal reference.

Another model which could be labeled semantic is described in Hobbs (1986). Unlike Heim and Kamp, Hobbs is explicity concerned with determining pronominal reference. Hobbs' 'semantic analyzer' however, goes beyond information available directly from the lexical items and structure of an utterance

---

[5]However, other work done individually by these two authors does concern local discourse relations and, as cited in the appropriate section, served as the precursors to the theory that I will use to treat null arguments.

and assumes that world knowledge "is available in the form of predicate calculus axioms. ...several 'semantic operations' ... draw inferences selectively from the collection of axioms." (1986:345-346) In other words, though Hobbs states that the input to his semantic analyzer is a logical representation of the text, this representation includes non-linguistic information.

Hobbs begins by assuming that syntactic conditions on non-co-reference have already applied. Four 'semantic' operations are then carried out on the input. First any of a small set of inter-sentential connective patterns are searched for (eg. contrast, cause, violated expectation, temporal succession, paraphrase, parallel, and example). If one is found then he predicts that certain interpretations will be strongly preferred.

Hobbs' second and third steps deal with the properties of predicates. In the second operation, he looks at 'predicate interpretation', i.e, what pronoun interpretations might be preferred based on the properties of a particular predicate. His third operation is 'knitting'; two predicates are determined to be identical if they are directly related through some information available from the interpretation of certain terms. This is roughly equivalent to one predicate being inferable from the other (cf. Prince 1981). When two predicates are considered to be identical, they are 'knitted' together, thus identifying the antecedent.

The final operation, used to resolve any remaining untreated NPs, is to 'identify entities' . This step handles the introduction of new discourse entities for indefinites, etc... Also, for any remaining unresolved 3rd person pronouns, "... a backward search through the Lexicon is conducted for a chain of inference that begins at some statement in the text and ends with a known property of the pronoun." (1986:348) If more than one plausible antecedent is in this last step, then shared saliency properties are used to determine the preferred interpretation.

In between syntactic and semantic approaches are theories such as that in Hajičová and Vrbová (1982). They argue that there is a hierarchy of salience and that this hierarchy is the most important factor in determining reference. They consider a mixture of semantic and structural/pragmatic factors to be

important for determining the hierarchy. Specifically, they argue that the important factors are whether an entity is part of the focus or the topic of an utterance, and whether it is referred to by a weak pronoun (or deleted). The specific rules are that 1) if an entity is expressed by a weak pronoun or a deleted form then it maintains the same level of saliency, 2) if it is represented as an NP in the focus then it gets the maximum level of salience, 3) if it is a definite NP in the topic, then it get the next highest level of salience, 4) if some entity referred to in an utterance gets an increase in saliency then all objects associated with this object (inferables) get a level of saliency two levels less than that entity, 5) if an object a is used in some expression like "As for a" or "Concerning a", then it gets one less then the maximum level of salience (i.e, the same level of salience as definite NPs in topics) , and finally, 6) if an object is not expressed in or associated with any entities represented in the utterance then it goes down two levels in salience. They argue that a difference of just one degree in salience (i.e topic vs focus) is not enough for disambiguating subesequent references.

Another mixed model is described in Polanyi (1988). Although this 'Linguistic Discourse Model' is not specifically a model of local discourse, it is a bottom-up model, building global discourse out of a variety of types of local units (segments?). Polanyi (1988) does not explore reference resolution but suggests that this model could be helpful. The model is mostly concerned with discourse SEMANTICS -- semantic building blocks are used to model hierarchical structures -- however, I categorize it as mixed for two reasons. First, the basic units are clauses and 'connectors' rather than semantic entities such as propositions. Second, many of the contextual factors associated with more complex discourse structures in this model depend on metalinguistic information such as particular types of speech events, plans, jokes, etc... Again, since pronoun resolution is an inferential process that involves non-linguistic information as well as linguistic competence, the incorporation of this information may have practical value, but it is not the ideal environment for describing the linguistic impacts of grammatical or referential options.

Like Polanyi (1988), Webber (1990) describes a model which is not concerned solely with local discourse but which includes some aspects of local discourse in its bottom-up approach. Webber is

primarily concerned in this paper with the resolution of the referents of demonstrative pronouns. She argues that discourse deixis can refer to discourse segments that are nodes on the right frontier of a discourse tree structure (which, as in Polanyi (1988), is a record of the process by which a discourse grows). She states, "... it is the structure of discourse segments (and hence that of the discourse model) that constrains the referents of deictic pronouns rather than the world being described." (1990:18) Unlike Heim and Kamp, Webber is concerned with the effect of discourse structure, including the ordering of propostions and their relation to each other, on what discourse entities are in the focus space (i.e. on the right frontier).

Each of the works described here has strengths and weaknesses that are a result of the general theoretical approach taken. I have tried to outline the most significant of these strengths and weaknesses here. However, each of these works also has strengths and/or weaknesses in the actual models developed within their theoretical frameworks. I have not discussed these details because, though they are certainly of interest, they are not actually relevant for choosing between the theories in principle.

5.3. CENTERING THEORY

The theory that I proposed to use to treat the properties of null arguments in English is CENTERING. Centering theory is concerned exclusively with the effect of clearly linguistic information on local discourse structure. Therefore, it is both more interesting theoretically and more easily implementable than Hobb's semantic theory. More specifically, Centering is concerned with how entities in the 'focus space' are ordered with respect to each other, and how this affects the coherence of discourse. It treats the structure of discourse segments in more detail than Polanyi (1988) or Webber (1990), and is therefore more useful for the treatment of null arguments. As will be shown, it is not unlike Hajičová and Vrbová (1982) but it takes a simpler approach to relevant features than their work because it specifically limits itself to discourse segments and does not rely on numerical assessments of relative

saliency. Centering also has the advantage over this theory of having been used to treat null arguments in a variety of other languages.

5.3.1. Comparing Centering Theory to DRT

As mentioned above, Heim and Kamp have extended the idea of the domain of formal linguistics beyond the level of a single sentence by pointing out that the definition of certain semantic operators cannot be explained in a sentence grammar. They've also concentrated on a method for representing what a pronoun *can* mean. In this respect, they handle some things which centering theory has not yet looked at, but they've also missed some of the constraints that centering does already include. In particular, they make the naive assumption that, if its representation is not in the scope of some operator that says otherwise, an entity is available as an antecedent to pronouns forever. Centering, on the other hand, helps in the resolution of pronominal reference (or at least in the reduction of ambiguity beyond DRT) by providing a preference ranking for interpretations.

In other words, centering picks up where Heim and Kamp leave off in two crucial respects -- 1. To the extent that both these theories provide a method for constraining what entities are available as possible antecedents to a pronoun, centering gives an explanation for preferences in pronoun interpretation, and 2. Though the list of available antecedents suggested by centering does not provided an account of scope effects and should probably either incorporate the work of Heim and Kamp or use the ouput of DRT as a filter, it does provide constraints not accounted for by them. Specifically, it allows entities to be available as antecedents for pronouns only if they are 'present' (in some roughly defined sense) in the previous sentence. The two theoretical approaches can therefore be thought of as complimentary.

5.3.2. Some Basics and an Existing Centering Model

The notion of a discourse center was introduced in Joshi and Kuhn (1979) and Joshi and Weinstein (1981). The ideas were influenced by work on discourse focus, such as that in Grosz (1977), and Sidner (1981, 1983). A broad centering theory is outlined in Grosz, Joshi and Weinstein (1986) and the first formally implemented version is found in Brennan, Friedman and Pollard (1987).

Grosz, Joshi and Weinstein avoided any strict pragmatic/cognitive definition of the notion center to avoid getting involved in the murky conflict over what 'topic', 'focus', 'theme', and other such terms refer to. Therefore, for the time being , we will simply say that the 'center' is the most salient/necessary/important thread of continuity in a discourse segment.

Because of this role that centers play in the discourse segment (i.e., the domain of local discourse), the center of a particular utterance is determined, in general, by looking back at the previous utterance and comparing it to the current one. We call a center which is determined in this way a 'backward-looking center' (Cb).

Of course, the process of continuity must somehow be started at the beginning of the discourse. As there is no previous utterance to look back to, a different technique is used. The identity of the center in an initial utterance is treated as a variable and is determined after its possible points of continuity with the subsequent utterance are available. We establish the center of the first utterance at that point using the hierarchy of the forward-looking centers list (defined below.) The center determined in this way will henceforth be called the 'established center' (Ce). (cf. Walker, Iida and Cote 1994)

The basic intuition behind centering theory is that certain structural features of an utterance serve as guides to the attentional structure of the discourse segment of which the utterance is a part. In particular, a speaker will design her utterance (U) not only to include a point of continuity with the previous utterance (U-1) but also to indicate how likely it is that each entity represented in U will be the

crucial point of continuity (i.e., the backward-looking center (Cb)) in the next utterance (U+1). A hearer, in turn, may choose between ambiguous interpretations of an utterance by considering these features. The (partially?) ordered set that results from ranking entities as possible future centers has been called the *forward-looking centers (Cf-) list*.

Loosely speaking, the intuition here is that a hearer will assume that, if possible, a speaker will be most likely to have the same Cb associated with two consecutive utterances in a discourse segment, and will also most likely construct an utterance in such a way as to make the Cb the most highly ranked entity on the Cf-list (i.e, the preferred Center (Cp)).

Brennan, Friedman and Pollard (1987), building on the notions of transitions in Grosz, Joshi and Weinstein (1986), suggested that the four combinations possible from these two properties result in four transition types between utterances. CONTINUATIONS, the cases where both properties hold, are the most highly preferred transitions. RETAINS, where the first property holds but not the second, are next. Third are the transitions where the Cb has changed but the new Cb is the Cp; these are SMOOTH-SHIFTS. Last and least preferred are ROUGH-SHIFTS; neither property holds in these transitions.[6] The predictions of this ordering are that the more highly ranked the transition between two utterances, the more coherent and easy-to-process is the discourse at that point, and also that, in the case of multiple possible interpretations for an utterance, the interpretation(s) that represent the most highly ranked transition available are the most preferred interpretation(s).[7]

A functioning model of centering must have an ordered set of features $f_1 ... f_i$, for a particular language l, which map to the particular entities E, in a particular utterance U, to determine the Cf-list for that utterance. Thus, the process of creating a Cf-list can be expressed as the following rule:

---

[6]Brennan, Friedman, and Pollard (1987) actually called the two types of shifts shift-1 and shift, respectively.

[7]It should be noted that this leaves aside the issue of the interaction of centering transitions with situational and world knowledge biases. Again, centering accounts for the influence of extra-sentential linguistic information on salience and pronoun resolution.

(1). **Cf-List Generation**

$$\text{Cf-list}(U_i^j) = E(e_1, ..., e_i) \text{ s.t. } \forall e((e_x > e_y) \Rightarrow (f^j(e_x) \geq f^j(e_y))).$$

To avoid confusion, I will henceforth call this ordered set of features for a language the CF-TEMPLATE for that language. In other words, I am making a terminological distinction between the general set of ranked features for a language, and the actual set of ranked entities in a particular utterance, which I will continue to call the Cf-list.

The exact nature of $f_1...f_i$ for English (and other languages) is not yet resolved. Ultimately, centering theory will work as a model of attentional state for a particular language only if we select the appropriate features in that language for the Cf-template. There are a priori any number of syntactic, semantic and lexical characteristics that could affect the ordering of Cfs. Grosz, Joshi, and Weinstein (1986) suggested that many different features may be significant. Though they left the issue relatively open, they did propose that the syntactic structure of an utterance (especially subject position) and the use of pronominal forms are important factors. In fact, an important corrollary to the rule ordering transitions is that the no other entity in an utterance will be pronominalized unless the Cb is pronominalized.

Brennan, Friedman and Pollard (1987) subsequently demonstrated, when they implemented the first formal model of centering, that many English examples can be handled using just grammatical relations in the Cf-template. In other words, they used the following Cf-template:

(2). SUBJ > OBJ > OBJ2 > OTHERS

However, no assumption was made that these were the the only features needed for the English template, nor even that they were the best possible features.

Of course, it would be highly unrealistic to think that a mechanism of this sort is the sole solution to determining the antecedent of a pronoun or null argument, but it would be equally unrealistic to ignore the structural aspect of discourse, of which centering is a part. Inferential factors such as world

knowledge, prejudices, gender biases, etc... all play an additional disambiguation role for the interlocuters. But these factors are not clearly linguistic and are best considered separately, perhaps after we have a better understanding of syntactic/semantic structural effects on local discourse.

5.3.3. The Current Approach to Ranking Forward-Looking Centers in English

The role of grammatical relations in linguistic theories has never been completely resolved, even at the level of sentential grammar. There has been no consensus on whether grammatical relations are linguistic primitives or simply surface features which are derivable from syntactic configurations and often pattern in predictable ways with thematic relations. In theories such as Lexical Functional Grammar (cf. Bresnan 1978), grammatical relations are indeed treated as linguistic primitives, but in Government and Binding Theory they are not primitives and, when they are referred to, it is simply as a shorthand for positions in a tree structure. (cf. Chomsky 1981, Haegeman 1991.) If it turns out that grammatical relations are not syntactic primitives than this could affect whether we should use them as primitives of discourse models (although this is not strictly entailed...) On the other hand, even if they *are* primitive concepts, they are not necessarily the best one for capturing the full range of linguistic facts relevant to centering transitions.

Nonetheless, as described above, the only existing model for centering in English has relied exclusively on grammatical relations. This Cf-template is actually sufficient to disambiguate many cases of third person pronominal reference.

5.3.4. Some Simple Applications of the Existing Centering Model

Below is an example of a simple discourse segment structure where a series of CONTINUATIONS represent the preferred interpretations for the utterances. In each of these utterances MARJORIE is both the Cb and the Cp.

(3). U_n: Marjorie works in New York.

Cb:? Cf: [$MARJ_s$ > $NEW\ YORK_{o2}$]

U_{n+1} : She shares an office with Anne.

Cb: MARJ Cf: [$MARJ_s$ > $ANNE_{o2}$] = continue

U_{n+2} : She always beats her to work.

Cb: MARJ Cf: [$MARJ_s$ > $ANNE_o$ > $WORK_{o2}$] = continue

Example (4) illustrates that the highest discourse transition available can sometimes be a RETAIN, as in U_{n+1} , or SMOOTH-SHIFT, as in U_{n+2} .

(4). U_n: Marjorie works in New York.

Cb:? Cf: [$MARJ_s$ > $NEW\ YORK_{o2}$]

U_{n+1} : Anne shares an office with her.

Cb: MARJ Cf: [$ANNE_s$ > $MARJ_{o2}$] = retain

U_{n+2} : She always beats her to work.

Cb: ANNE Cf: [$ANNE_s$ > $MARJ_o$ > $WORK_{o2}$] = smooth-shift

The RETAIN is forced in U_{n+2} because ANNE is in a more highly ranked position (subject) than MARJORIE and hence must be the Cp. Consequently, since the existence of two feminine pronouns in U_{n+2} requires that both ANNE and MARJORIE be realized in that utterance, the subject gets interpreted as ANNE in that utterance as well and the transition is a SMOOTH-SHIFT.

5.3.5. Limitations of the Existing Model

The nature of the Cf-template is one of the most crucial issues in Centering theory. The use of grammatical relations has served as a useful starting point and has worked well for the certain cases, but there are obvious limitations to this approach.

Issues in languages other than English have already forced centering researchers to begin augmenting cf-templates with other relevant features besides grammatical relations. For instance, Kameyama (1985, 1986) demonstrated that, in Japanese, topic markers (-wa) on NPs and empathy markers on verbs also have strong effects on how highly ranked an entity is. In her analysis, these factors were simply added as additional components of the Cf-template for Japanese. Walker, Iida, and Cote (1990, 1994) demonstrated that including these features in the Japanese Cf-template was relevant for the full range of centering transitions and that topic effects were sometimes important for the determination of preferred interpretations even in utterances with zero pronominals (i.e., without an overtly wa-marked topic.) Di Eugenio (1990) argued that features on the verb, as well as functional considerations, can effect interpretations (and hence entity-ranking) in Italian. Rambow (1992) has argued that word order is relevant to German Cf-templates in subordinate clauses. Most recently, Gordon, Grosz and Gilliom (1993) have argued that word order (especially in the case of non-argument positions in modifer clauses expressing the perspective of one of the salient entities) can effect the English Cf-template as well.[8]

There are also a number of well-recognized types of reference in English which are problematic for a purely syntactic approach to Cf-templates. These include deictic and event reference, 'inferables' (cf. Prince 1981), and temporal references. In short, it is quite clear that Cf-templates can not be composed entirely of grammatical relations nor even, more generally of just syntactic features.

---

[8]Thus far none of the features, other than those discussed for Japanese, have been incorporated into specific formulations of Cf-templates for these languages.

Most importantly, from the perspective of this thesis, grammatical relations do not provide an environment for representing the discourse effects of implicit objects in English. As shown in Chapter 3, null objects in English are not present in the syntactic representation of sentences. Therefore, to incorporate them into a centering model, it is necessary to make lexical information available to the Cf-template.

In the next section, I describe a revised model in which the Cf-template is based on the primitives of Lexical Conceptual Structure, thereby providing a means for incorporating null arguments, and null objects in particular, into the model. I show that this revision does not hurt the ability of the model to account for other cases, and does perhaps suggest a means for treating other problematic data such as deictics, discourse deixis and temporal reference.

5.4. REVISING CENTERING TO HANDLE NULL ARGUMENTS

Any alternative that is considered should at the very least, incorporate all the work already done using grammatical relations and, in addition, provide a method for handling at least some of the currently untreated cases. Ideally, it should also suggest a reason for the relative effectiveness of grammatical relations. Perhaps most important of all, at least from the perspective of a linguist, it should share with grammatical relations the property of concerning itself with *linguistic* features, rather than with features of the more general and largely unmanageable realm of knowledge representation. In other words, it should explore the limits to which linguistics can provide an account of local discourse structure.

In the remainder of this chapter, I will discuss evidence from implicit null arguments in English and suggest that the use of lexical conceptual structures (cf. Jackendoff 1987, 1990, 1993) is a viable alternative that meets the above criteria.

5.4.1. An LCS Centering Model

In chapter 3, I proposed that lexical conceptual structures (LCSs) as defined in Jackendoff (1987, 1990, 1993) offer the best framework for representing the grammatical analysis of null objects in English. I now propose that structures of this general type are also useful for ranking entities in local discourse structure. For convenience, LCSs for the SOA verb *enter* and the IOA verb *eat* are repeated below.

(5). LCS for enter (Jackendoff 1990:46):

$[_{Event}$ GO $([_{Thing}]_i, [_{Path}$ TO $([_{Place}$ IN $([_{Thing}]_j)])])]$

(6). LCS for eat (Jackendoff 1990:253, slightly modified):

[CAUSE $([_{Thing}]^{\alpha}_i, [_{Event}$ GO $([_{Thing}]_{<j>}, [_{Path}$TO $([_{Place}$ IN

$([_{Thing}$ MOUTH-OF $([\alpha])])])])])]$[9]

(7). LCS for enter when the object has been lexically affected (i.e., is null):

$[_{Event}$ GO $([_{Thing}]_i, [_{Path}$ TO $([_{Place}$ IN $([_{Thing}$ $\mathbf{0}_1])])])]$

(8). LCS for eat when the object has been lexically affected (i.e., is null) :

[CAUSE $([_{Thing}]^{\alpha}_i, [_{Event}$ GO $([_{Thing}\mathbf{0}]$, $[_{Path}$TO $([_{Place}$ IN

$([_{Thing}$ MOUTH-OF $([\alpha])])])])])]$

By substituting the LCSs of particular lexical items of the correct type in for the variable elements of verbal LCSs like these, we can build more complex, phrasal, conceptual structures. This follows from the principle of lexicalization described by Jackendoff (cf. Jackendoff 1983:185) and repeated below:

---

[9]It is debatable whether the 'eat' event is really a special case of GO. Dorr 1990 treats eating as a primitive. The distinction is not, however, important for the purposes of this paper.

(9). **Jackendoff's Lexical Variable Principle**

A variable in the structure of a lexical item must be capable of being filled by a conceptual constituent.

More specifically, the features of the argument are combined ("fused") with those already specified by the verb in a unification-like process (cf. Shieber 1986) which Jackendoff (1990) defines as follows:

(10). **Argument Fusion** (Jackendoff 1990:53)

To form the conceptual structure for a syntactic Phrase XP headed by a lexical item H:

a. Into each indexed constituent in H's LCS, fuse the conceptual structure of that phrase YP that satisfies the coindexed position in H's subcategorization feature.

b. If H is a verb, fuse the conceptual structure of the subject into the constituent indexed *i* in H's LCS.

Example (11) shows the fused structures of two sentences with the verb enter.

(11). a. The dog entered the room.

$[_{Event}$ GO $([_{Thing}$ DOG], $[_{Path}$ TO $([_{Place}$ IN $([_{Thing}$ ROOM])])])]

(Jackendoff 1987:183)

b. The dog entered.

$[_{Event}$ GO $([_{Thing}$ DOG], $[_{Path}$ TO $([_{Place}$ IN $([_{Thing}$ $\mathbf{0}_1$])])])]

(modified version of above as defined in Chapter 3)

If the phrasal conceptual structure of an utterance is the only place where crucial information (such as the presence of a null object in (11)b) is represented, then this structure must somehow be able to provide input into the centering process. I suggest that the Cf-template should refer directly to these

structures and their components, rather than to grammatical relations, as major components of the Cf-template.

This suggestion is not as unusual as it might seem at first. In fact, using θ-roles would simply be a less powerful approximation of the conceptual structure approach. As Jackendoff (1990:47) says, "...thematic roles are nothing but particular structural configurations in conceptual structure; the names for them are just convenient mnemonics for particularly prominent configurations." With a cf-template for English based on θ-roles, the highest θ-role represented in an utterance would be ranked most highly on the cf-list for that utterance; the next highest role would be next on the list, and so forth. For simple cases, this would be equivalent to ranking the subject highest, and so forth, as is generally done now. A Cf-template with lexical conceptual structures can therefore be thought of as subsuming (rather than replacing) what is already represented by currently used factors.

Jackendoff argues further that although the correct representation of thematic roles can only come from a good understanding of conceptual structure; it may be that the θ-structure (the part of CS relevant to syntactic structure) of a lexical item is only a small subset of its full conceptual structure. Though this may be the case, it would be inappropriate to assume a priori that the same is true for the relationship between local discourse structure and conceptual structure. Though θ-roles alone could express what is now expressed with grammatical relations, some or all of the additional information in conceptual structures may also be relevant to attentional state. For instance, using θ-roles (as they are currently understood) as the source of the Cf-template's elements would provide no way for the Cf-template to refer to the whole event or state represented by an utterance, but using conceptual structures, which explicitly refer to these events and states, would, in contrast, provide direct access to this information. Since both pronouns and null subjects and objects in English may have event antecedents, this is an important addition.

The crucial question is: How do we form the Cf-template, i.e., how do we rank the components of a conceptual structure? Previous work on Cf-templates composed of grammatical relations is relevant to this question.

The relation between θ-roles and conceptual structures has already been mentioned. The thematic hierarchy has received a great deal of attention in the GB literature. Using this hierarchy, thematic relations could be substituted into existing centering analyses with little problem; therefore corresponding conceptual structure components should be transferable as well.

In an LCS, the hierarchical relationships between the elements is represented by their position in the structure, i.e, the left-to-right ordering goes roughly from more highly ranked elements to less highly ranked elements (see Jackendoff 1990:258.)[10] In general then, statements made about grammatical relations can readily be mapped into relationships between the elements of the LCS of a verb (eg., the special status of 'subjects' would correlate to the special status of the leftmost entity in such an LCS).

Assuming the essential truth of this parallel (a more in-depth study of the hierarchical relationship between the various primitive elements of LCS, and especially of how these hierarchies merge in more complex derived conceptual structures is certainly warranted), an alternative Cf-template for English can be defined using the following rule:

(12). **Lexical Conceptual Structure Cf-template for English (first approximation):**

Rank arguments of the phrasal conceptual structure for U_n from left-to-right. [11]

---

[10]Actually, Jackendoff describes a more complex method which involves action tiers and thematic tiers, but a details of this discussion are not critical for our purposes here.

[11]This first definition could alternatively be expressed as a partially-ordered list of the (external) arguments of conceptual structure operators (Like CAUSE, GO, FROM, etc...) based on how they combine, such as:

CAUSER > GO-ER/MOVER/... > BE-ER > FROM-SOURCE, TO-GOAL/... > ...

(i.e., a list roughly based on correspondences to the thematic hierarchy)

When we use conceptual structures, we gain the extra descriptive power described above. Even the first approximation of the Cf-template for English shown in (12) makes accessible one of the advantages of conceptual structures -- it provides a means for the discourse entities represented by null objects to be included in the ranking, while still capturing the relationships between overt arguments as well. A straightforward use of the LCS in (8) above provides a Cf-list for (13), which explains the availability of the null object in U_{n+1} for subsequent reference in U_{n+2}.

(13). artificial example

U_n : Mary invited us to her new restaurant.

U_{n+1} : We ate all night. = retain

CF-list (a) : [CAUSER(we) > GO-ER(*thing eaten*) > MOUTH-OF(we)]

U_{n+2} : It was all delicious. = smooth-shift

Notice that 'we' occurs twice in this ranking (as the CAUSER and as the argument to MOUTH-OF), but only the highest occurrence is relevant. This is not a surprising state of affairs; work on Japanese centering in utterances where a discourse entity was both the TOPIC and the OBJECT showed that only the most highly ranked feature was relevant for ranking in these cases as well (cf. Kameyama (1985, 1986) and Walker et al. (1990, 1994)).

A better Cf-template would also lower the ranking of null objects, as in (14) below.

(14). **Lexical Conceptual Structure Cf-template for English (revised)**:

Begin by ranking arguments of the phrasal conceptual structure from left-to-right.

All 0_i arguments in U_n are equal in ranking to the lowest ranking overt arguments in U_n .

This revision has the effect that arguments affected in the lexicon and indexed to an existing discourse entity will be less preferred as Cbs of the next utterance (as argued in Chapter 3), though they will still be allowed. The new definition explains the (temporary) ambiguity in (15).

(15). [Hopper (1992:149)]

A_1: We **ate** at-at Jorges

Cf: [CAUSER(we) > GO-ER(**0**)/ MOUTH-OF(we)

/PLACE (jorges)[12]

B_1: Was it goo:d?

A_2: Um:: it was all ri:ght

B_2: pt -hhhh you know, I don't think that's as good as a lot of people think it i:s

The GO-ER (the thing eaten) in $A_1$is not ranked more highly than the PLACE entity because it is a null object, i.e. lexically-affected. Therefore, it in B_1 (and also A_2) is as likely to refer to one of these as to the other. Only the fact that the speaker then clearly refers to 'Jorges' with a deictic (that) makes it less likely that the PLACE was also the intended referent in the previous utterances.

5.4.2. The Status of Transitions

Not all the effects of null objects are accounted for with this one change however. The effect of the minimal constrast between the second utterances in (16) and (17) on the interpretation of the two identical third utterances requires an additional observation.

[12]This last entity would be taken from the more complex conceptual structure incorporating the LCS for the preposition.

(16). a. Harry was feeling lonely.

Cb: ?Cf: [HARRY]

b. Bill called him.

Cb: HARRY Cf: [BILL, HARRY] (retain)

c. He talked for two hours.

Cb: HARRY Cf: [HARRY] (continue)

(17). a. Harry was feeling lonely.

Cb: ? Cf: [HARRY]

b. Bill called.

Cb: HARRY Cf: [BILL, HARRY] (retain)

c. He talked for two hours.

Cb_1: BILL Cf_2: [BILL] (smooth-shift)

Cb_2: HARRY Cf_2: [HARRY] (continue)

In these two (artificial) examples, the only difference is that Harry is syntactically present (and overt) in the first and only present in the conceptual structure of the second. The result is a preferred interpretation of (16)c with Harry doing the talking, but two more or less equally preferred interpretations of (17)c (i.e., either Harry or Bill could be doing the talking). This does not follow immediately from our revised Cf-template for English because, in each case, Bill is already ranked more highly than Harry. There is a CONTINUE transition available for the third utterance of both examples and this corresponds to the same interpretation of each utterance. Thus, some other part of the centering mechanism must be involved.

I believe the answer lies in the ranking of transitions. Though it does seem to be the default for CONTINUES over other transitions, we can easily find situations, independent of null object contexts, in which the opposite needs to be true. For instance, in (18)d the word back is a cue to the hearer to choose the RETAIN transition for this utterance.

(18). a. Bob hasn't been getting along with Ed lately.

b. He saw him on the street last night

Cb: BOB Cf: [BOB, ED, STREET]

c. and he punched him in the arm.

Cb: BOB Cf: [BOB, ED, ARM] (continue)

d. Then he hit him back.

Cb_1: BOB Cf_1: [ED, BOB] (retain)

??Cb_2: BOB Cf_2: [BOB, ED] (continue)

Clearly, the default preferential ranking of transitions can be rearranged for a particular instance. The word back can be thought of as an operator on this ranking and, I believe, so can the use of null objects. In other words, the relative likelihood of different centering transitions occurring at any point in a discourse segment can vary and null objects are one of the factors that can determine the relative status of each transition.

5.4.3. Summary

I have argued here that by using the ordered elements of Lexical Conceptual Structures as input to the Cf-template of a centering model, we can continue to rank correctly those discourse entities already captured by the use of grammatical relations, while at the same time allowing for the availablility and effects of implicit lexical information like null objects. I claim that these implicit arguments have an impact on local discourse coherence which was previously unincorporated, and that the properties of these arguments also offer additional evidence that the preferred ordering of centering transitions is somewhat more complex than was first proposed.

5.5. ADDITIONAL BENEFITS OF THE REVISED MODEL

In the early centering model, where discourse entites were ranked based primarily on grammatical roles, adding additional features to the Cf-template sometimes involved mixing apparently different types of information. For example, the Walker et al. centering analysis of Japanese zero pronouns mixed the syntactic information of grammatical roles with the more pragmatically-influenced features of TOPIC and EMPATHY. [13] (I hesitate to call them strictly pragmatic features since they are grammatically marked in this 'discourse-oriented' language.)

Once we have accepted that lexical-conceptual information effects the ranking of discourse entities, we have a more elegant environment for including other 'conceptual' information about the discourse. In other words, we can talk about what otherwise seem like a variety of different factors equally, as conceptual elements or types, and use the same mechanism for entering them into the discourse model.

In discourse-oriented languages, this may include grammaticalized pragmatic information like topic and empathy. Even in English, the structure of a discourse may be influenced by reference to events described within or entailed by the discourse, but not previously referred to.

A shown in Chapter 3, the presence of EVENTS as entities in LCSs is crucial to representing the constraints on arbitrary null objects. It seems that, in addition, a Cf-template built of lexical conceptual elements can handle the availability of entities realized as null event reference as antecedents to pronominal references. Suppose that a simplified LCS for <u>see</u> is something like (19) below:

(19). LCS for <u>see</u> when the sentential complement is lexically-affected:

$[_{Event}\ \text{SEE}\ ([_{Thing}\ ]_i, [_{Event}\ \mathbf{0}_1])]$

[13]Thanks to Susumu Kuno (private conversation) for pointing out this unexplained conflation.

The null event reference is now directly represented in the input to the Cf-list for (20)B_2, making it available as an antecedent to that in (20)B_3.

(20). [revised version of Hopper (1992:69) (utterance B_3 added)]

A_1: I'm the one that goes to UT that got you all the fun information

B_1: oh really?

A_2: yeah hhhh

B_2: Oh I *see*.

[Event SEE ([Thing]i, [**Event** **0**1])]

B_3: well, *that's* quite a coincidence.

More generally, *any* event represented in a conceptual structure, not just those in argument positions, may be available for pronominal reference.[14] For instance, in the example below, the pronoun that in the third utterance refers not to what B sees but to the whole event of B's seeing.

(21). A: I don't think we should go because it's very expensive.

B: Ok, I see.

[**Event** SEE ([Thing]i, [Event **0**1])]

A: Well, that's good because I was afraid you'd be upset.

Because this event is also present in the conceptual structure of the sentence, it is accessible to the Cf-list, though additional data studies are needed to determine how highly ranked they should be.

5.6. OPEN ISSUES

One interesting aspect of a conceptual structure analysis of cf-templates is that it seems to have some potential common ground with the idea, discussed in Grosz, Joshi and Weinstein 1986, that

[14]Though this may be subject to other constraints not relevant here.

centering theory should fit into semantic *situations*. The availability of event reference is one area of overlap that is already clear but there could, in principle, be others as well.

The interaction of deictic reference with other types of pronominal reference is an issue which ultimately must be incorporated into any model of attentional state. The simplest possible resolution for centering is to place discourse participants on every Cf-list, ranked equally with the most highly ranked entity from each utterance. However, the identity of these discourse participants must come from somewhere. Thus far, all the elements of derived conceptual structures has come from the fusing of lexical structures. We could however, also associate semantic concepts with the uttering of a phrase clause -- elements identifying the speaker, hearer (s), and the time and location of the utterance. Though this approach as roughly described here is not strictly identical with the idea of situation semantics referred to by Grosz, Joshi and Weinstein (based directly on the ideas in Barwise and Perry 1983), there are clear parallels that would be worth exploring.

I believe it would also be interesting to attempt to handle so-called 'inferables' using this revised centering model. It is unclear which types of inferables, if any, can be represented with the limited types and amounts of conceptual information that are part of lexical or phrasal *structure*. The implications of a split along these lines among types of inferables is that they vary in their impact on discourse structure. More specifically, we might predict that those inferables which can be derived from phrasal conceptual structures will be more like null objects in this respect, while those that can not (eg. those that rely on conceptual relationships not central to lexical structure) will have different or lesser effects. There is in fact, some evidence of such a split.

In example (22)b below, the pronoun refers to an entity (the class 'dalmation') which has not been explicitly evoked in the discourse but which was inferable from the explicit reference to a member of that class in (22)a (cf. Webber (1986)).

(22). a. The judge awarded "Best of Show" to *a dalmation* this year.

b. I have always believed that *they* are great dogs.

Like subsequent pronominal reference to null objects, this pronoun is perfectly felicitous. Both of these cases contrast with the one illustrated in example (23) below, in which a pronominal reference to an inferable is not acceptable:

(23). a. I went out to my car on the street.

b. The windowshield/ #it was shattered.

It is possible that the distinct here is due to the fact that, though the existence of a windshield is pragmatically inferable from the existence of a car given our knowledge of what cars look like, windshields are not included in the lexical conceptual structure of cars. In contrast, the conceptual structure of a dalmation is built from the LCS of dalmation and the class dalmation is implicit in much the same way as a null object.[15]

Finally, whether or not the notion of conceptual structure described here would work as well in building the Cf-templates for languages other than English must be considered an open issue. One promising piece of information is that lexical conceptual structures have been used with a great deal of success in at least one machine translation project (cf. Dorr (1990)), but the features relevant to local discourse structure in other languages still need to be considered directly.

[15]The one significant difference is that the class entity is not a variable that needs to be affected in the lexicon in order to be absent from syntactic structure.