

Chapter 5

Conclusions

In this chapter I will summarize my findings, sketch a preliminary processing model for the comprehension of reference, and suggest directions for future work.

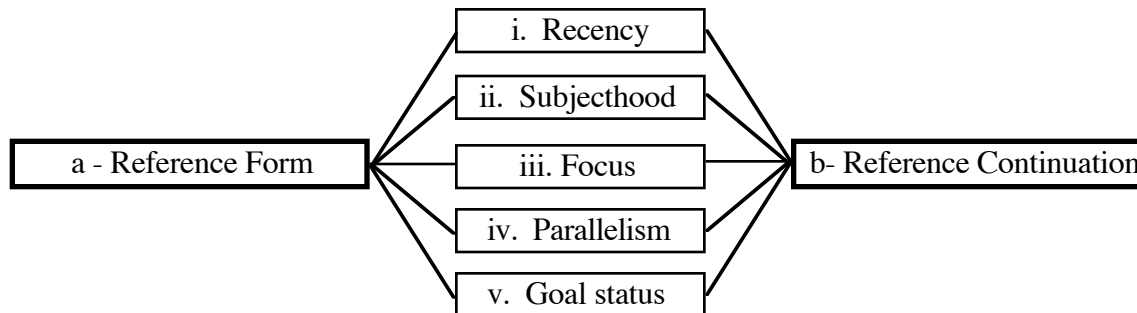
5.1 The findings

The goal of this dissertation was to understand why diverse linguistic factors influence choices in reference form. I considered five factors: Recency, Subjecthood, Focus, Parallelism, and Goal Status. All five had previously been shown to influence reference processing, yet there is otherwise nothing obvious that binds them together.

Through text analyses, corpus analyses, and experimentation, I demonstrated that all five factors share another property: each is associated with referents that have a higher probability of being continued in the following discourse than other comparable referents. Recently mentioned referents are more likely to be referred to than referents that have not been mentioned for awhile; subject-referents are more likely to be subsequently mentioned than object-referents; focused referents more likely than nonfocused referents; parallel referents more likely than nonparallel referents; and goal-referents more likely than source-

referents. A summary of these findings is depicted in Figure 2 from chapter 1, repeated here as Figure 1.

**Figure 1: A summary of the findings in this dissertation:
All five factors are associated with both Reference Form and
Reference Continuation**



These data lead to the generalization that people use attenuated forms, like pronouns, to refer to the types of things they usually talk about.

There are two main questions that arise from these data. First, why do speakers tend to refer to referents associated with these five factors? I have suggested that for each factor, patterns of reference are best understood under the "language-as-action" approach to studying language use (H. Clark, 1996). This approach focuses on the fact that people use language for a purpose, whether that purpose is to coordinate a business transaction, relate a narrative, give a lecture, or something else. This means that discourses are driven by the goals and intentions of the discourse participants, at multiple levels. I have argued that the goal-driven nature of language underlies the association of Reference Continuation with each of the factors in Figure 1.

The second question that arises is why speakers use less-specified forms for referents with these properties. The language-as-action approach again offers an answer: people speak with the intention of communicating with their addressees (H. Clark, 1996). This means that on any given occasion, they design their utterances for their addressees, given particular local and global aims. From this generalization it is plausible to assume that

speakers choose referential forms that will facilitate comprehension. Less-specified forms of reference are more efficient for all discourse participants, but in some cases comprehenders need more information to correctly identify the referent. This type of "listener accommodation" explanation also appears in other accounts of reference form and reference processing (e.g., Ariel, 1990:16; Clark and Marshall, 1981).

Because my account is one of listener accommodation, it is important to understand how each factor relates to processes of comprehension. In the preceding chapters, I suggested that the association of each factor with Reference Continuation has implications for how they are used during language processing. In the following section, I will offer some further speculations about how this information might be stored and used.

5.2. A constraint-based account of activation

Recent models of psycholinguistic processing have suggested that learning a language involves learning the correlations between forms and functions, and the degree to which they correlate, at multiple levels (e.g., Bates and MacWhinney, 1989; MacDonald et al., 1994; Tanenhaus and Trueswell, 1995; Trueswell et al., 1994). Once learned, this information is used during language comprehension.¹

In this section, I will review two such approaches to language comprehension. Then I will propose a similar model for reference processing.

5.2.1. Competitive, constraint-based models

The Competition Model is one example of a competitive, constraint-based model that involves the use of probabilistic, correlational information (Bates and MacWhinney, 1989). Bates and MacWhinney argued that, during language comprehension, people interpret the input on the basis of competing "cues", which differ from one language to another. For example, understanding any sentence requires identifying which NP is the grammatical subject. In English, word order is a fairly reliable indicator of grammatical

APPENDICES

functions, so English speakers generally take the pre-verbal NP to be the grammatical subject. In contrast, word order in Italian is more flexible than in English, so it is less informative. (Bates, MacWhinney, and Smith 1983; Bates 1976). Instead, Italians rely more on pragmatic and morphological cues in identifying subject NPs. Bates and MacWhinney measured the relative strength of cues in different languages by presenting people with unusual or ungrammatical sentences and forcing them to decide on an interpretation. For example, when English speakers were presented with a sentence like "The pencil hits the cow", they tended to choose "the pencil" as the agent/subject. In contrast, when Italian speakers were presented with the Italian analogue "la mattia colpisce la vacca", they tended to choose the cow ("la vacca"). They claimed that Italians do not "trust" word order to the degree that English speakers do, and rely instead on pragmatic cues during the assignment of grammatical functions (Bates and MacWhinney, 1989:11).

Bates and MacWhinney argued that cross-linguistic variation in language comprehension can be studied in terms of the cues that are relevant for each language. The strength of a cue for a particular language can be measured in terms of its "cue validity", or the reliability of cues in a particular language (1989:41). While cue validity is a characteristic of the language itself, its psycholinguistic corollary is "cue strength":

This is a quintessentially connectionist notion, referring to the probability or weight that the organism attaches to a given piece of information relative to some goal or meaning with which it is associated. In other words, cue strength is the weight on the connection between two units.

In our psycholinguistic instantiation of this idea, each link between a given surface form and an underlying function is given a weight or strength. With this kind of mechanism, no sharp line is drawn between probabilistic

¹ Theoretically, this information should be used in production as well. However, most studies in these traditions so far have focused on comprehension processes.

tendencies and deterministic rules. An obligatory relationship between form and function is nothing other than a connection whose strength approaches unity. This permits us to capture statistical differences between adult speakers of different languages, and it permits us to describe facts about language change (in language history and/or in language learning) in gradual and probabilistic terms; we are not forced to postulate a series of all-or-none decisions in which parameters are definitively set and rules are added or dropped. (Bates and MacWhinney, 1989:42)

Thus, cue strength is the weight of a given piece of linguistic information. For example, word order has a higher cue strength for English speakers than for Italian speakers for decisions about grammatical functions. According to this model, language acquisition is essentially learning the relative strengths of cues for a target language.

A similar approach has been developed in constraint-based models for the processing of lexical and syntactic ambiguities (e.g., MacDonald, 1993, 1994, 1996; MacDonald et al., 1994; Tabor et al., 1998; Spivey-Knowlton and Sedivy, 1995; Trueswell, Tanenhaus, and Kello, 1993; Trueswell, Tanenhaus, and Garnsey, 1994). These models propose that readers or listeners interpret ambiguous input on the basis of all available constraints. Constraints, like Bates and MacWhinney's cues, are features of the input (linguistic or otherwise) that support one interpretation over another. When all constraints point to the same interpretation, comprehension is easy, but when they conflict, comprehension is hard. Furthermore, these models suggest that syntactic and lexical ambiguities are processed with the same general mechanisms.

For example, a much-studied syntactic ambiguity is the main verb/ reduced relative ambiguity, as in 1.

APPENDICES

- (1) a. The horse raced past the barn.
b. The horse raced past the barn fell. (Bever, 1970)

MacDonald et al. (1994) and others have argued that syntactic ambiguities like this one derive from the lexical ambiguity of "raced". In 1a, "raced" is a past tense main verb, while in 1b it is a past participle verb in a reduced relative clause. However, the two sentences appear identical until the very end, creating a temporary structural ambiguity. MacDonald et al. propose that ambiguous words like "raced" have ambiguities at multiple levels of representation, and that biases at each level can support one interpretation or another. For the main verb/ reduced relative ambiguity, they discuss five levels of ambiguity, listed in Table 1 (MacDonald et al., 1994:689)

Table 1. Levels of ambiguity for the Main Verb/ Reduced Relative ambiguity

- a. syntactic structure (main verb vs. reduced relative)
- b. tense morphology ("ed" as past tense vs. "ed" as participle)
- c. voice (active vs. passive)
- d. argument structure (transitive vs. intransitive)
- e. assignments of thematic roles to nouns in the sentence (e.g., subject as agent vs. subject as theme)

For example, consider the fragment "The evidence examined...." To arrive at a reduced-relative interpretation of "examined", it is also necessary to assign the role of theme to "the evidence". The process of assignment is facilitated here, because "the evidence" makes a good theme, of "examine" but not a good agent. The reduced relative interpretation also requires the comprehender to interpret "examined" in the transitive frame. This process is also facilitated, because "examine" is an obligatorily transitive verb. Thus, MacDonald et al.

argued that the reduced relative reading should be easier for fragments like this one than fragments without these properties.

Their claims have been supported by several studies. For example, Trueswell (1996) showed that the processing of this ambiguity is influenced by biases at two of these levels. First, it is influenced by the degree to which the initial noun is a plausible theme or agent for the following verb ("e" in Table 1). For example, it's plausible to think that a room would be searched, but not that it would search something, so "the room" is a good theme, but a bad agent. Therefore "the room searched" promotes the reduced relative interpretation, where the initial noun would have to be the theme. In contrast, "the thief" is a bad theme, but a good agent, so it would promote the main verb interpretation. Second, it is easier to arrive at a relative-clause interpretation if the "ed" verb form is often used as a participle, relative to the overall usage of the verb ("b" in Table 1).

In a word-by-word self-paced reading task, Trueswell had participants read sentences like those in 2 and 3.

- (2) Exp. 1 (Trueswell, 1996): Initial noun is good theme, bad agent
 - a. The room searched by the police contained the missing weapon.
 - b. The room that was searched by the police contained the missing weapon.

- (3) Exp. 2 (Trueswell, 1996): Initial noun is good agent, bad theme
 - a. The thief searched by the police had the missing weapon.
 - b. The thief that was searched by the police had the missing weapon.

The stimuli compared ambiguous reduced relative sentences (2a and 3a) with their unambiguous counterparts (2b and 3b). Half the items used verbs like "search" that had a low participle frequency (Low-PP items), half used verbs with a high participle frequency (High-PP items).

APPENDICES

The results were reported in terms of the difference between the ambiguous sentences and the unambiguous sentence, in different regions of the sentence. The crucial region is at "by the police", which provides unambiguous support for the reduced relative reading. In Experiment 1, where plausibility supported the relative clause interpretation, the results showed processing difficulty only for the Low-PP items. That is, when the "ed" verb form was frequently used as a past participle, it biased the interpretation towards the reduced relative sufficiently that no processing difficulty occurred. In Experiment 2, by contrast, plausibility supported the main clause interpretation, and reading was slowed for both Low-PP and High-PP items. But the verb bias played a role in this experiment as well, in that it took longer for readers to "recover" from the difficulty in the Low-PP condition than the High-PP condition. So, in the High-PP condition for Experiment 1, the reduced relative interpretation was supported by the plausibility information, the frequency information, as well as the disambiguating by-phrase. Since all constraints agreed with each other, no processing difficulty occurred. In other conditions, where the constraints did not support the same structure, processing was relatively slower. Thus, Trueswell's study demonstrated that syntactic ambiguity resolution is sensitive to contextual semantic constraints, as well as to the frequency with which a given form is used for a particular function.

Although the term "constraint-based" suggests that these models are principally about the use of multiple constraints, their most important feature is that the constraints are probabilistic and graded (Spivey-Knowlton and Tanenhaus, 1994:436). Trueswell et al. (1994) described their model of parsing as follows:

The principles that underlie the approach are simple. Structures are partially activated with the strength of activation dependent upon their likelihood given the input. The effects of a contextual constraint will depend upon its strength and the availability of the alternative structures.

(Trueswell, Tanenhaus, and Garnsey, 1994:304)

Thus, syntactic ambiguity resolution is influenced by essentially two things: the strength of the available constraints and the competition among alternative interpretations. Structures compete with each other in parallel, and are activated by constraints like lexical frequency and semantic plausibility.

For the purposes of my proposal, it is important to highlight the role that frequency plays in these models. Like the Competition model, constraint-based models assume that part of learning a language is learning the constraints of that language, and that constraints are built from a person's experience. One type of constraint is characterized in terms of lexical frequency information. In the example above, individual verbs differed in terms of the frequency with which they occurred in the form of a past participle relative to their overall frequency. The relative frequency of transitive and intransitive verb uses also influences the resolution of main verb/ reduced relative ambiguities (MacDonald, 1994). The relative frequency of uses for a lexical item is hypothesized to be stored with that lexical item, available for use during language comprehension.

A particularly important aspect of these models is their emphasis on the role of the individual's experience in learning generalizations about the world, and thus the degree of constraint associated with different types of information. In some cases, constraints are associated with purely linguistic information, like the relative frequency of participial uses for a given verb. In other cases, however, constraints come from experience with the real world and the plausibility of certain events. Therefore it is possible for non-linguistic information to influence linguistic processing.

5.2.2. Reference processing

My proposal here is to approach reference processing in terms of constraint-based models. These models have been developed to deal with ambiguities at the lexical and syntactic levels; here I will consider referent identification as a type of ambiguity resolution

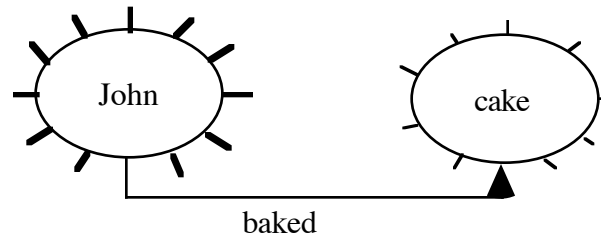
APPENDICES

in the referential domain. From these models I will adopt the features described above: representations can be partially activated, based on probabilistic information that supports one interpretation over another. Activation changes dynamically as linguistic input is processed, and as different sources of information become available.

There are numerous other assumptions that my proposal makes, based on other work in discourse processing. These assumptions were detailed in §1.2.2.2; I will briefly review them here.

First, I assume that discourse processing involves a representation of the information shared between interlocutors, or the "common ground" (H. Clark, 1996). One part of the common ground is the discourse representation, which contains both the text model and the situation model. The text model is where linguistic input is briefly represented and parsed before it is used to update the situation model. The situation model contains information about the linguistic and nonlinguistic aspects of the discourse. In a face-to-face conversation, for example, it would contain information about the physical surroundings, the identity of the interlocutors, and also the linguistic component of the discourse. In a written text, by comparison, the reader's model would contain primarily the information from the text. The situation model includes representations of the discourse entities, the events, and the relationships among them. The representation of a particular entity might best be considered as a distributed representation, consisting of the conglomeration of features for that entity (Kawamoto, 1988; Seidenberg and McClelland, 1989). For ease of presentation, however, I will use local representations for discourse entities and other information in the situation model.

Thus, the sentence "John baked a cake" might have a situation model like that in Figure 2. An important feature of the representations of entities is that they can be activated to a greater or lesser degree. Activation, as before, is represented as the number and size of lines emanating from a particular representation.

Figure 2: Partial situation model for "John baked a cake"

When comprehenders construct text and situation models, they also need to access information stored in long-term memory. Grammatical knowledge is needed to parse the linguistic input. For example, "John ate the cake", results in a different (and more plausible) model than "John was eaten by the cake". World knowledge also plays a role in constructing the situation model. For example, if Alfred hears Sharon say "I bought a chair", she might envision a prototypical chair, with four legs, a straight back, and no arms. But Sharon's mental model might be different if she knows that Alfred was in the market for a big, comfy armchair. World knowledge also serves as the basis for making inferences, which also influence the situation model. For example, the sequence in 4 would lead to the inference that the picnic supplies contained beer, which would then be represented in the situation model.

(4) We checked the picnic supplies. The beer was warm.

(Haviland and Clark, 1974)

The comprehender's knowledge in long-term memory is also important for determining the activation of entities in the situation model. I have shown that five factors are associated with an increased probability that their referents will be referred to in the following utterance. I propose that this information is stored in long term memory, along with other information about those structures. For example, Jane's experience with the

APPENDICES

English language has endowed her with a concept of grammatical categories. She knows that subjects and objects map in systematic ways onto the thematic roles of a verb. I propose that her representation of the subject category also represents her experience with the discourse role of subjects, and her experience that subject-referents have a high probability of being referred to in the subsequent discourse. This knowledge might be represented in one of several ways. It may be represented directly in terms of the probability of subsequent reference. Or it may be represented in terms of an index of the speaker's attention -- that is, Jane's experience is that speakers tend to talk about subject-referents; this tells her that subject-referents are usually important to speakers, so they have a high probability of subsequent reference. Regardless of how it is represented, Jane's knowledge of the discourse role of subjects results in higher activation for subject-referents in her situation model. Her representation of the focus category in cleft constructions has a similar property, since through experience she knows that the focus-referent of a cleft has a higher probability of subsequent reference than the nonfocus-referent.

For both subject and focus categories, the information associated with them can be said to form part of Jane's linguistic competence. Similarly, her experience with parallel structures is also part of her linguistic knowledge. Jane's experience with discourse processing has shown her that parallel structures often have parallel referents. There are several possibilities for how information about parallelism may be represented. It may be stored in terms of syntactic structures, as in Figure 3.

Figure 3: A possible partial representation of parallelism information

Given two subsequent clauses with this structure:

$NP_1 [VP NP_2]$

$NP_i [VP NP_j]$

There is a high probability that $i=1$ and $j=2$.

Parallelism may also have a more complex representation, such that the more features shared by two verbs, the higher the likelihood that the referents are parallel (Arnold, 1995, Smyth and Chambers, 1994).

In addition to linguistic knowledge, Jane also has world knowledge that influences her situation model and discourse processing. One relevant aspect is her knowledge that recency of mention is correlated with probability of subsequent reference. This might be represented in terms of the amount of time that has passed since a referent was mentioned, or the amount of new material that has been introduced to the discourse. On the other hand, this knowledge might be represented more generally in terms of focus of attention: the more recently something was mentioned, the more likely it is to be in the focus of attention of Jane's interlocutors. Finally, Jane's experience with the world has also taught her that people tend to be interested in goals and consequences. This information is also available to Jane when she constructs a situation model of a goal referent.

In sum, people's experience with a specific language and the world around them is recorded in long-term memory. I propose that part of their knowledge includes their experience of discourse patterns. Thus, for each of the five factors I have discussed, people represent their association with Reference Continuation.

How does this knowledge influence discourse comprehension? In the following section I will make some suggestions.

5.2.3. Referent activation during discourse comprehension

The following is a preliminary sketch of how referent activation may work during discourse comprehension. For presentational purposes, I will adopt a variant of the Interactive Activation model to describe my suggestions for referent activation (McClelland and Rumelhart, 1981). This and similar models were originally proposed to account for word recognition. MacDonald et al. (1994) used this architecture to describe a unified model of lexical and syntactic ambiguity resolution, in which they proposed to include syntactic information in the lexicon. So previous uses of the model have restricted themselves to processes relevant to the text model. Here I am proposing to extend the model one step further, to account for representations of entities in the situation model. This step is a reasonable one to make, considering the evidence that parsing is influenced by the referential context of an utterance (e.g., Tanenhaus et al., 1996).

The interactive activation model has several properties helpful for describing my data. First, it works in terms of an activation metaphor. It consists of representations at several levels, each of which is called a "node". Nodes are local representations at each level, which means that each node represents the whole of a particular representation. For example, McClelland and Rumelhart (1981) posit a node for the word "cart", as well as nodes for the letters "c", "a", "r", and "t", and nodes for each letter feature (e.g., a horizontal line). There are advantages to using distributed representations instead, as some later connectionist models do (e.g., Kawamoto, 1988; Seidenberg and McClelland, 1989). However, I will stick with local representations for simplicity.

Nodes are connected to other nodes both within the same level and at adjacent levels, such that the activation of one node influences the activation of connected nodes. Nodes can be connected with either an excitatory or an inhibitory connection: excitatory connections increase the activation of connected nodes, inhibitory connections decrease it. In addition, the connections between nodes vary in their strength, which is specified as the "weight" between two nodes. The weight between two particular nodes reflects the degree

to which one node supports the activation of another. In the case of an ambiguous word like "bat", the word node would be connected to semantic representations for both "bat-as-flying-mammal" and "baseball bat". The use of "bat" for the "baseball bat" meaning is more common than the "flying mammal" meaning (at least for some speakers of American English), so the weight connecting the word "bat" to the meaning "baseball bat" would be relatively stronger.

At each stage of processing, each node can be activated to varying degrees, depending on the input. When an activated node stops receiving activation, it decays until it reverts to its initial state. This initial state, or "resting activation", may be different for different nodes. An important feature of the interactive activation model is that activation on a given node builds up over time. Thus, it contrasts with other models of language processing that assume that accessing information is essentially instantaneous (McClelland and Rumelhart, 1981:379).

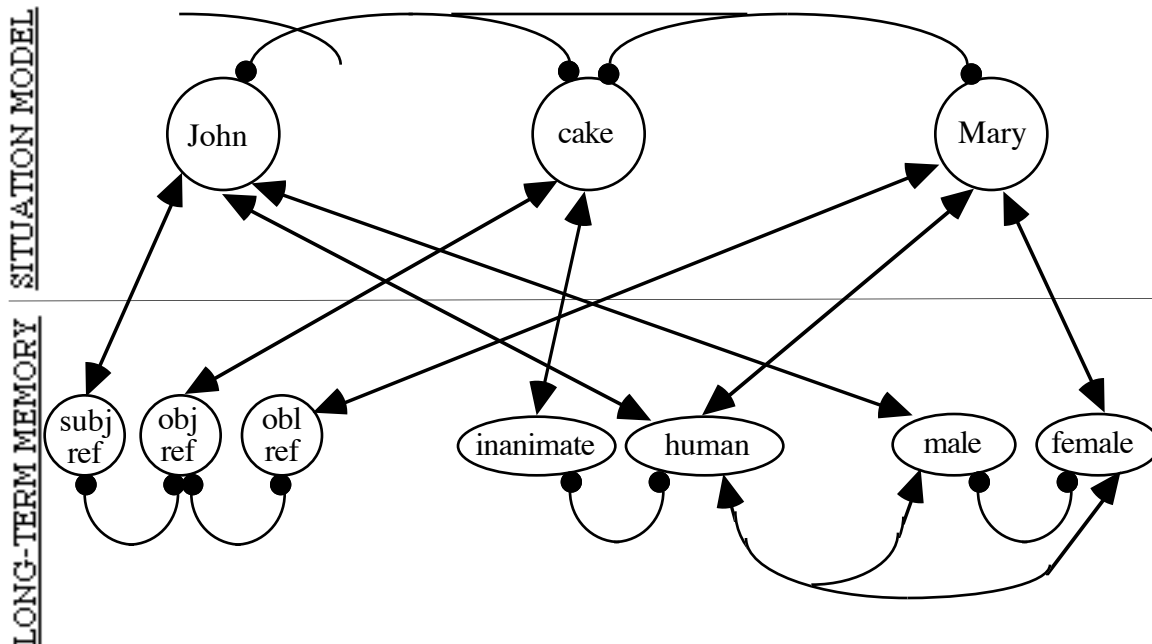
For example, consider a situation in which Jane hears the sequence in 5. Assume that the characters in this sequence are known to both Jane and the speaker.

- (5) a. John baked a cake for Mary.
- b. The next day he gave it to her.

When Jane hears 5a, she parses it and instantiates representations of "John", "cake", and "Mary" as nodes in her situation model. These nodes are associated with two kinds of nodes in long-term memory. First, they are associated with her conceptual knowledge of the referents. For individuals like "John" and "Mary", she accesses representations of those people; for indefinite referents like "cake", she accesses a representation of that type of object, which is influenced by her experience with prototypical instances of that category, and the most plausible instance for that given situation. Second, the entity nodes are linked with properties of the linguistic phrases used to introduce them to the discourse. Note that these properties are not a part of the situation model per se; rather, the discourse entities are

linked with the comprehender's stored knowledge about linguistic categories and their pragmatic uses. A partial picture of these representations is shown in Figure 4. In this picture, the only aspect of the linguistic phrase I am including is the information about grammatical function. Following McClelland and Rumelhart (1981) I am representing excitatory connections with arrows, and inhibitory connections with circles.

Figure 4: A partial model of the entities in "John baked a cake for Mary"



At the next time step, Jane parses 5b, which contains the pronoun "he". At this point, Jane is faced with the task of identifying the referent of the pronoun. Since pronouns are almost always used for referents that are already in the situation model of the listener, Jane knows to look for the referent there.

I propose that anaphor resolution proceeds in the following way: First, the referent-seeking anaphor becomes connected to properties in long-term memory. These connections are weighted to the degree that the listener knows that the referent for the anaphor has those properties. For example, given the anaphor "he", Jane knows that the referent is probably

male and human. Therefore, the weight of the connection between "he" and the nodes for "male" and "human" is relatively high. At the same time, the anaphor develops inhibitory connections to the properties "female" and "nonhuman". Note that these connections represent the process of searching for a referent, and not the process of identifying the anaphoric expression itself.

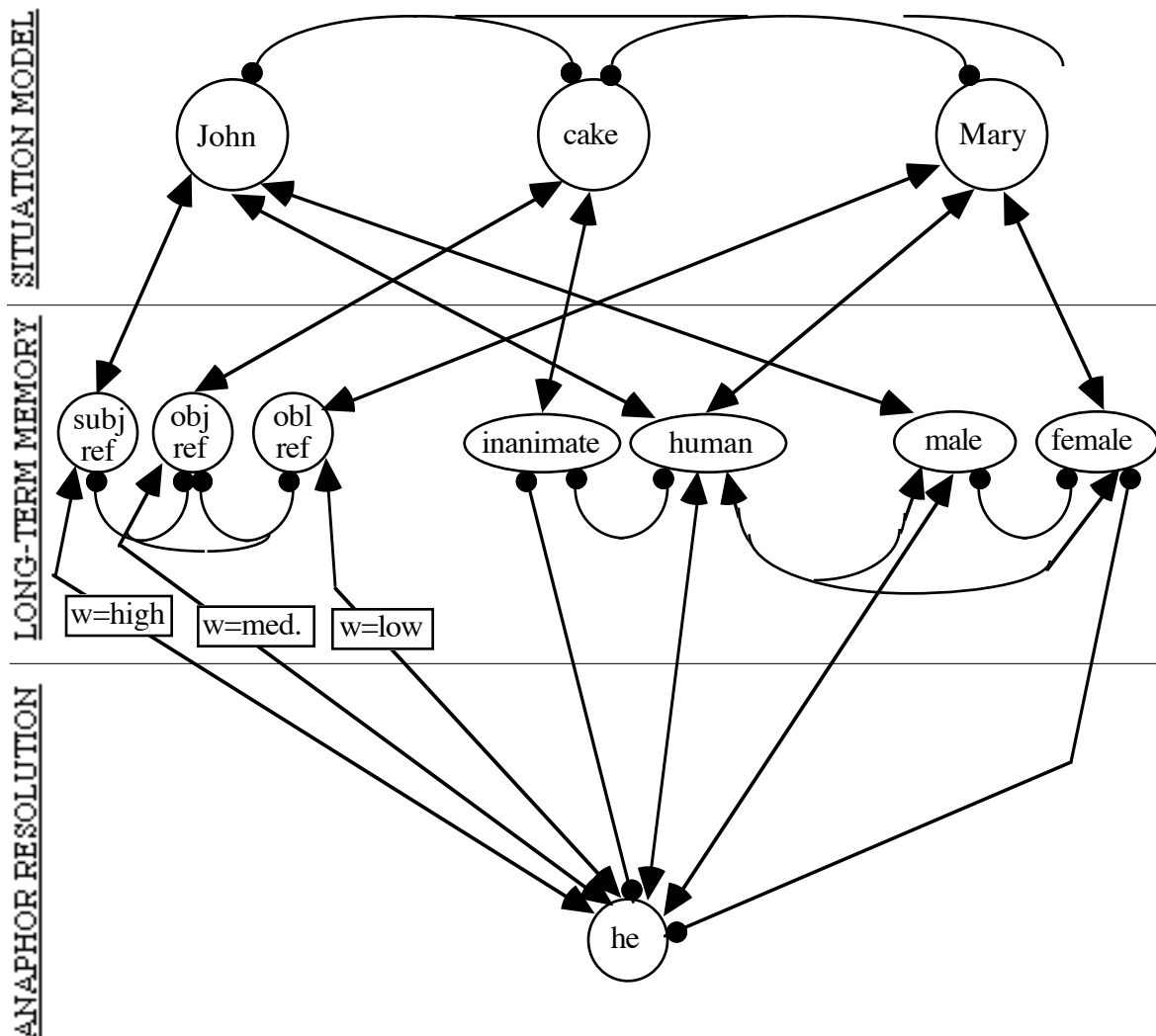
In addition to properties like "male" and "human", the anaphor becomes connected to properties of the referents in the discourse model. These properties include information from each of the five factors I have discussed. In the preceding chapters I showed that each factor is associated with "Reference Continuation". I proposed that this information is stored as a part of an individual's linguistic knowledge, such that each factor carries information about the probability of subsequent reference. Here, I propose that this information is encoded in terms of the weights of the connections. In this example, Jane has the knowledge that the referents of anaphors are more likely to be subject-referents than object-referents or oblique-referents. Therefore, the weight for the connection with the node for "subject-referent" is higher than that for "object-referent", which is higher than that for "oblique-referent". In an actual implementation of this model, the weights would be specified as real numbers between 0 and 1. Here, however, I have merely indicated the **relative** weights of the connections, with the labels "high", "medium", or "low".

A portion of the referent resolution process is depicted in Figure 5. This picture is necessarily only a partial one, and makes one particularly dangerous oversimplification. In past chapters I have talked about referent activation partly in terms of how the referent is instantiated into the situation model. That is, subject-referents are instantiated with higher activation than object-referents. This view is consistent with the idea of "salience", that some referents have inherent properties that make them more activated. In this figure, however, I am focusing on the point in time where the anaphor is encountered. At this point, the relative probability of each referent is especially relevant, because the processing system is engaged in the task of identifying a referent for the anaphor. By placing the weights on

APPENDICES

the connections between the anaphor and the properties "subject-referent", "object-referent", etc., I am representing the fact that this information is relevant to the process of anaphor resolution. However, entities in the situation model do not start out with the same levels of activation, and factors like Subjecthood can influence the activation of referents from the moment they are entered into the situation model. This kind of activation is discussed further in §5.2.4, below.

Figure 5: A partial model of referent resolution for the segment "John baked a cake for Mary. The next day he..."



Consider what happens as Jane tries to identify a referent for "he". The node for the anaphor partially activates the nodes for the properties of "subject-referent", "object-referent", and "oblique-referent". The only one that is connected with "John" is the node for subject-referent, so the partial activation associated with the subject-referent accrues to "John". Note that if "John" had been the object-referent, the same process would have occurred, except that the weight for the object-referent node is lower, so "John" would have received less activation. At the same time, the node for "he" sends activation to the nodes for "male" and "human". These connections are weighted relatively high, because the referents of "he" are usually male and human. Since "John" is also connected with these nodes, activation flows along these path to the node for "John".

While activation for "John" rises, the other referents also momentarily experience low levels of activation. "Mary" and "cake" receive low levels of activation from their association with the properties "object-referent" and "oblique-referent". The node for "Mary" is also associated with the "human" node, so it receives some activation from this route. However, the "female" and "male" nodes are mutually exclusive, so they inhibit each other. Therefore, as "male" rises in activation, "female" falls, thus decreasing the activation for "Mary". Similarly, the nodes for "John", "Mary", and "cake" are mutually inhibited, so as one node rises in activation, the other two will fall. This ensures that only one referent will "win out" as the correct choice. In this example, the unique match between "he" and "John" is sufficient for John to be chosen as the referent, but this kind of categorical cue is not necessary for successful referent activation, since information from the linguistic expression and inferential processes are hypothesized to also affect referent activation. An important feature of the interactive-activation model is that it allows the "wrong" choices to become partially activated. However, low levels of activation are not noticeable to the individual.

Thus, this example illustrates how Subjecthood affects the comprehension of an anaphoric reference. The support for "John" as the referent of "he" comes from three

APPENDICES

sources: 1) his association with the feature "male", 2) his association with the feature "human", and 3) his identification as subject-referent at the current moment in the discourse. It is important to note that the activation from his identification with the feature "subject-referent" is only partial. By itself, it would not be enough to fully activate the node for "John". In combination with other information, however, it increases the rate at which the "John" node becomes activated, thus facilitating the interpretation of the anaphor.

The role of the Subjecthood constraint becomes clearer if we compare the previous example with the variant in 6.

- (6) a. John baked a cake for Mary.
- b. The next day she thanked him for it with a bottle of wine.

As with the previous case, the anaphor node ("she") is connected with the nodes for subject-referent, object-referent, and oblique-referent. These routes send partial activation to both "John" and "Mary", but relatively more activation to "John" as the subject-referent. At the same time, the anaphor "she" has excitatory connections with the properties "female" and "human". This ultimately results in the selection of "Mary" as the referent, but full activation occurs somewhat more slowly than in the previous example.

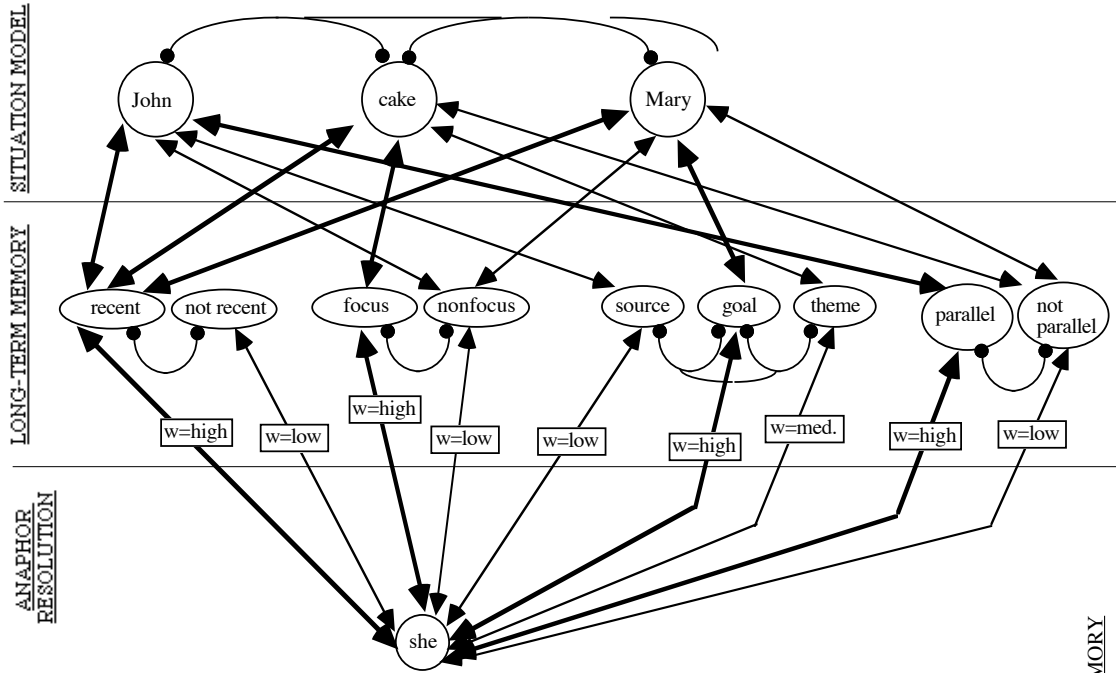
Just as this model can account for the influences of the Subjecthood constraint, it can also account for the other four factors discussed here. In each case, the bias toward one anaphor type over another is encoded in the weights of the connections. As an illustration of the remaining factors, consider the processing of the sequence in 7.

- (7) a. What John gave to Mary was a cake.
- b. She was really surprised.

Figure 6 shows a partial representation of the factors relevant for processing the anaphor "she" in 7b. The bold lines indicate the routes of activation for the connections with "high"

weights, so the reader can better appreciate the effect of each factor on the activation of referents.

Figure 6: A partial model of referent resolution for the segment "What John gave Mary was a cake. She..."



In sum, this model encodes the effect of each factor in terms of the weights on the connections between the anaphor and each value for a given factor. The weights themselves can be thought of as the probability of output y given input x . That is, speakers of English learn over time that subject-referents have a high probability of being referred to in the following clause. Therefore, given an actual anaphoric form, there is a relatively high probability that the referent is the subject-referent of the previous clause. The competing values for a particular factor (e.g., subject-referent vs. object-referent) are also mutually inhibitory. This means that the actual activation of a given node is the combination of the excitatory activation and inhibitory activation from other nodes.

5.2.4. The dynamic nature of activation: pre- and post-ambiguity constraints

Language comprehension is an inherently dynamic process. It is not always possible to build a fully-specified interpretation of the input immediately, but evidence shows that processing is immediate and incremental for processing at lexical (Marslen-Wilson, 1990), syntactic (Frazier, 1987), and referential levels (Tanenhaus et al., 1996). This also means that the listener's model at any given point is a function of the information that is available at that point. During the interpretation of lexical and syntactic ambiguities, it has been noted that comprehension can be influenced by information that comes **before** the ambiguity itself (pre-ambiguity constraints), as well as information that becomes available **after** the ambiguity (post-ambiguity constraints) (MacDonald, 1994, 1996).

Recall that I am treating reference resolution as a type of ambiguity resolution. Therefore, information that is available before the referring form itself can be considered a type of "pre-ambiguity" constraint. Consider the case where Jane hears "John baked a cake for Mary...." The subject-referent, John, has a higher probability of being subsequently referred to than either "cake" or "Mary". At the same time, there is a high probability that the discourse will continue after this clause, and furthermore that the following discourse will relate to the current utterance in some way, perhaps with an anaphoric reference. In

APPENDICES

essence, Jane can expect her interlocutors to be coherent. Therefore, Jane can **anticipate** the occurrence of an anaphoric reference or inferentially related information to some degree. In anticipation of some unknown anaphoric reference, and based on the information that is available at this point, her representations for each discourse entity become partially activated. That is, the discourse referents receive partial activation before the anaphor itself is encountered. This can be considered "predictive" activation: subject-referents are more predictable than object-referents, so they become relatively more activated. Another way of saying this is that subject-referents have a high probability of being related to the speaker's local goals and intentions, and thus are relatively topical. However, it is important to realize that this level of predictability is partial, and it occurs as a side-effect of the aim of trying to process the input with efficiency.

It is also important to note that predictive activation will not have a noticeable effect in all cases. The information from the anaphor itself is far more important to its interpretation, so that no amount of predictive activation can override it. For example, even if you thoroughly expect a reference to "Mary", you won't mistake the word "John" as a reference to her. This reflects what Marslen-Wilson termed "bottom-up priority" (1987, cited by MacDonald et al., 1994).

Activation from pre-ambiguity constraints is only noticeable under two conditions. One is when an extremely strong pre-ambiguity constraint provides enough activation to facilitate or inhibit the processing of a subsequent anaphor. Two candidates for this kind of constraint are Recency and Subjecthood, which are both large, robust effects. Therefore, recent-referents and subject-referents have higher activation from the point they are instantiated in the situation model, and not just when an anaphor is encountered.

In chapter 3 I suggested that subjects and foci are tools for negotiating the salience of referents in common ground. Under the view of Subjecthood and Focus as pre-ambiguity constraints, "salience" can be considered to be the activation of referents in anticipation of a possible future reference to them. From the listener's point of view, this is

the same thing as building a situation model that closely matches that of the speaker: some referents are more activated for the speaker; these are the referents that the speaker intends to talk about or interact with non-linguistically (see §5.3.2.3, below). The listener uses cues like these to infer which referents have a high probability of being important in the following discourse. Put another way, listeners infer which referents are important for the speaker.

A second condition where pre-ambiguity constraints would be noticeable is when two or more weak constraints combine with each other. An example of this type is Parallelism. As I suggested in chapter 2, parallelism information may be useful in conjunction with other constraints that herald the coming of an anaphor. If a subject anaphor is imminent, subject-referents may receive partial activation. If an object anaphor is imminent (signaled by a transitive-biased verb), object-referents may receive partial activation. If there is not enough information to signal the coming of an anaphor, parallelism effects will not come into play until the anaphor itself is encountered. Another example of this type is Goal Status, which interacts with other constraints that signal the type of continuation. When other constraints suggest that the incoming clause specifies the consequences of the preceding event, the goal-referent receives partial activation.

At the other end of the scale, referent processing is also influenced by information that **follows** the anaphor. In many cases, this information results from inferential processes, based on knowledge of the world and the situation. Consider the example in 8.

- (8) a. Steve tried calling Tom. But he couldn't get through.
 b. Steve tried calling Tom. But he wasn't at home.

In both 8a and 8b, the underlined pronouns are ambiguous at the point they are encountered. There are two discourse entities that match the features of the pronoun, so neither one is immediately selected. At this point the nodes for "Steve" and "Tom" are partially activated on the basis of their associations with other constraints, and "Steve" as

APPENDICES

subject-referent would be slightly more activated than "Tom". It isn't until the following context is processed that the referent becomes clear. In 8a, the bias toward "Steve" would be confirmed by the following context, and processing would be facilitated, but in 8b the bias toward "Steve" would slightly inhibit the selection of "Tom" as the referent of the anaphor.

In sum, I have suggested that referent resolution is a dynamic process, where referent representations become probabilistically activated by means of supporting constraints. Referent representations also compete with each other, such that in most cases a given anaphor selects one and only one referent.

5.3. Future Work

The model I sketched in the preceding section is preliminary and somewhat speculative, and leaves many questions open. In this section, I will suggest a number of directions for future work. First, I will discuss how these data might be considered in an alternate model, Discourse Representation Theory. Second, I will discuss some specific questions raised by the approach I have taken.

5.3.1. Discourse Representation Theory

A useful direction for future work is to consider how my data might be treated with different theoretical architectures. The framework I chose for my discussion was that of constraint-based models. However, it is notable that these models have been developed to account for the processing of lexical and syntactic ambiguities. My extension of these models has the advantage that it uses the same mechanisms for discourse processing, and thus moves toward a unified account of language processing at all levels. However, it is worth considering how existing theories of discourse might account for the same data, and whether it would be profitable to adopt characteristics of other models. Although I must leave this topic for future research, I will briefly consider one such model here: Discourse Representation Theory (DRT).

DRT, (Kamp, 1981; Kamp and Reyle, 1993), is one example of a semantically-oriented discourse theory. It was originally developed to account for different kinds of anaphoric and scope relations, most notably the "donkey sentences", as in 9.

(9) If Pedro owns a donkey he beats it.

Every farmer who owns a donkey beats it. (Geach, 1962)

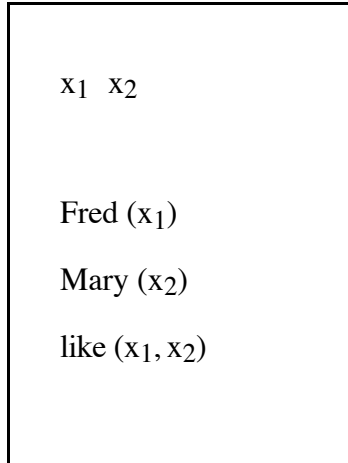
DRT is similar to other semantic theories in that it discusses information in terms of truth conditions, but one difference is that it looks at truth conditions over the domain of the discourse, rather than a sentence.¹

Under DRT, the representation of a discourse proceeds in terms of a Discourse Representation Structure (DRS). These structures contain representations of the discourse entities and their relations. Consider the example sentence "Fred likes Mary". First, the input is parsed into its syntactic structure. On the basis of this parse, the discourse entities are entered into the DRS. In this example, there are two discourse entities corresponding to "Fred" and "Mary". The discourse markers themselves are essentially variables, but they are given content by association with the referents denoted by each NP. Propositions are also specified in terms of discourse markers (DMs), as in 10.

¹ This review of DRT comes mostly from a summary presented by Stirling (1993:156-170)

APPENDICES

(10) The DRS from the sentence "Fred likes Mary" (Stirling, 1993:161)



Given this discourse structure, anaphoric references are linked to their referents by entering them into the DRS as DMs, and linking the DMs together with the notation $x_i = x_j$. An example of a larger discourse is given in 11.

- (11) The DRS from the discourse "Fred likes Mary. He gave her a cat. It is large and cunning." (Stirling, 1993:161-2)

x_1	x_2	x_3	x_4	x_5	x_6
Fred (x_1)					
Mary (x_2)					
like (x_1, x_2)					
cat (x_4)					
give (x_3, x_6, x_4)					
$x_3 = x_1$					
$x_6 = x_2$					
large (x_5)					
$x_5 = x_4$					
cunning (x_6)					
$x_6 = x_4$					

One advantage to DRT is that it treats anaphoric and nonanaphoric NPs similarly: in both cases, a DM is entered into the DRS, and the only difference is whether it is linked to another DM. However, using DRT to model my data raises several challenges, since it makes several assumptions that are in direct conflict with what is known about language processing. To reconcile DRT with my data and other psycholinguistic findings, a number of changes would have to be made to DRT.

APPENDICES

The first problem is that DRT builds the discourse model on the basis of the syntactic parse, thus positing two separate stages for syntactic and discourse processing. However, it is well established that syntactic processing proceeds incrementally, and that parsing decisions are immediately influenced by the referential context (e.g., Tanenhaus et al., 1996). This would not be possible if the parse were needed before the discourse representation could be built. Thus, it would be necessary to implement a procedure whereby syntactic and referential processing could proceed simultaneously and incrementally, and influence each other.

A second problem is that DRT seems to allow for only two psychological states for a discourse entity: an entity can either be available for anaphoric reference, or not. In contrast, my data and other findings on reference processing have shown that referents can have differing levels of activation in the discourse representation. Thus, DRT would need some way of representing differences in activation. One possibility would be to rank the DMs relative to each other, as Centering Theory does (Gordon and Hendrick, in press). However, this approach poses two problems. First, it requires that as the discourse proceeds, the relative ranking of the entities must change. However, it is not clear when the re-ordering process should take place. Centering ranks entities at the end of each sentence, but this does not capture the incremental nature of processing. The second problem is that my data do not support ranking as a mechanism for controlling activation. For example, the data in §4.3 showed that people used pronouns more to refer to object-of-PP-referents when they were goal referents than when they were source-referents. That is, goal / object-of-PP-referents were more activated than source / object-of-PP-referents. However, in both cases the subject-referent was relatively more activated than the object-of-PP referent. This means that a ranking mechanism would not distinguish between goal- and source-referents, because both would merely be ranked second, relative to the subject-referent.

More generally speaking, it would be necessary to modify DRT to allow it to account for dynamic processes of language comprehension. It might be possible to update

the DRS at smaller increments, like after each word. This would approximate a continuous model of processing, although it would not achieve the same level of continuity that is present in real language processing. For example, even lexical access proceeds incrementally, such that listeners can often identify a word before hearing all of it (Marslen-Wilson, 1990; Tanenhaus et al., 1996).

Finally, it would be necessary to develop some mechanism for generating the link between a given anaphor and referent. As DRT stands, links are merely posited. That is, it is possible for the linguist to represent the intended interpretation of an anaphor, but the model does not specify which referent is the most natural, or which processes would drive the choice of a referent.

In sum, using DRT to model my data poses a number of challenges. I have suggested a number of issues that would require attention if DRT were used to model psycholinguistic data, a task I leave for future work.

5.3.2. Further questions

Given the model I proposed in §5.2, a logical next step is to implement it, including all factors simultaneously. Although this would be a useful exercise, the model above makes many simplifying assumptions that would be problematic for a general model of reference processing.

5.3.2.1. Language processing as a dynamic process

One thing missing from the model above is a dynamic characterization of reference processing. Although the interactive activation model has a number of advantages for this kind of data, it represents a single point in time with respect to the input. Since McClelland and Rumelhart (1981) originally designed their model to account for word recognition in reading, it was reasonable for them to restrict the focus to a single time point, corresponding to an eye fixation on a word. Discourse processing, however, proceeds dynamically over time. So, for example, some constraints may come into play before an anaphoric form is encountered, whereas others may not be available until later.

5.3.2.2. *How continuous is Recency?*

The model in §5.2 also presents an overly simple view of recency as a dichotomous variable: recent vs. not-recent. However, the data in chapter 2 showed precisely the opposite, that recency influences reference form in a continuous way. At the very least, a model of reference processing needs to represent smaller distinctions in recency for referents, such as the number of clauses since the entity was first mentioned. One question for future research is whether recency is a completely continuous constraint, or whether it can best be characterized as discrete, on the level of the clause or a clause-segment. My data also show that recency effects are modulated by such factors as differential interference from main and subordinate clauses (§2.3.1.3). Further work is needed to understand how this difference interacts with recency in a processing model.

5.3.2.3. *Linguistic and nonlinguistic aspects of joint activities*

A third area that needs development is the relation between reference and other ways of maintaining coherence in a discourse. I have focused on the forms of anaphoric reference used to evoke previously-mentioned entities. However, speakers often use other forms of linguistic cohesion such as bridging inferences (Clark and Haviland, 1977; Haviland and Clark, 1974). In an example like 4, above, the interpretation of the second sentence requires the comprehender to access the representation of "picnic supplies" and integrate with it the representation of "beer". The same factors discussed here should influence such inferences: they should be easier if "picnic supplies" had appeared recently, in subject or focus position, or as the goal of the previous proposition.

By the same token, people's knowledge of Subjecthood, Focus, and other factors is built from their experience with both anaphoric reference and other inferential processes. That is, Mark's experience with English has shown him that when something is mentioned in subject position, it has a high probability of being important to the following discourse. I have discussed this mostly in terms of direct anaphoric references. However, I have assumed that the association between these factors and Reference Continuation

encompasses cases where the subsequent reference is indirect, as with inferences and part-whole relationships. For example, imagine that Jane says to Mark "The fruit at Safeway is really good this week. I especially like the oranges." Mark's comprehension of "oranges" requires accessing the previous referent "the fruit at Safeway". Even though this is not a direct anaphoric reference to a specific entity, Mark can still take it as evidence that the most topical referent of the first sentence was "fruit at Safeway", in the sense that this is what Jane intended to talk about. This example would count as an instance in which the subject-referent of the first sentence was continued in the next one, thus contributing to Mark's knowledge of the correlation between Subjecthood and Reference Continuation. I have assumed that Reference Continuation is sensitive to part-whole relationships and other types of anaphoric links; this is what motivated my decision to include categories "part of" and "related to" in the corpus analyses in chapters 3 and 4.

This idea can be extended even further. When Mark learns that the subject-referent is what is most likely to be important to the speaker in the following discourse, he knows more than just how the speaker will use that referent linguistically. More generally, he knows that the referent is important to the speaker, and will be important to their joint activity as a whole. I have focused here on linguistic factors and linguistic patterns, but language is often just one part of a joint activity, used along with non-linguistic types of action (H. Clark, 1996). Thus, once an entity has been brought into the common ground, participants in the joint activity can cause other participants to access it again, using either linguistic or non-linguistic means. For example, imagine that Jane and Mark are baking a cake together. If Jane hands a cake pan to Mark, she does not need to say anything, yet Mark will activate a representation for the pan. This could be called nonlinguistic reference -- Jane referred to the pan, but not with words.

Nonlinguistic reference probably plays a role in the formation of the categories of Recency, Subjecthood, Focus, and Goal Status. It would not play a role in Parallelism, because this depends on the structural relation between a linguistic anaphor and its linguistic

APPENDICES

antecedent. For the other factors, once a referent has been mentioned, especially as subject, as focus, or in the role of goal, it becomes more probable that it will be important to the following activity. This activity may include discussing it linguistically, manipulating an object physically, both. In all cases, the frequency with which it is important to the following activity is stored as a part of an individual's knowledge of that category (perhaps as a ratio of the number of times a particular type of referent is continued out of the total number of times that type of referent is encountered). Thus, Mark's experience with English and the world leads him to believe that things mentioned in subject position have a higher probability of being important to the speaker than things mentioned in object position. His experience comes from exemplars where the speaker said "Frank is late. He got stuck in traffic," as well as cases where the speaker said "The pizza's here!" and handed him a slice.

One direction for future exploration is the relation between linguistic factors like Subjecthood and the speaker's nonlinguistic actions. This poses numerous challenges, one of which is how to define "nonlinguistic reference". How do speakers make someone else pay attention to a particular object in the environment? Likely candidates are eye gaze or gestures (e.g., handing an object to someone). Another issue is the difference between things that attract attention for communicative reasons, and things that attract attention because of some inherent salience, like a loud explosion. If Jane and Mark are co-present during an explosion, Mark can assume that the event was activated for both of them, but he cannot assume any relation between the explosion and Jane's prior utterances. In contrast, if Jane says "the pizza's here!" and hands Mark a slice, he can reasonably infer a relation between the utterance and the subsequent gesture, and thereby learn facts about how people use language (see Baldwin and Tomasello, 1997).

5.3.2.4. Comprehension and production

I have discussed my general findings in terms of the activation of referents, focusing on how certain constraints are likely to influence reference **comprehension**. I have

assumed throughout that the comprehension of referring forms is important to the speaker's choices of form. I made this assumption on the basis of the observation that discourse is a joint activity, and therefore speakers design their utterances for the purpose of communicating with their interlocutors (H. Clark, 1996). I therefore assumed that speakers' choices are made in part to facilitate comprehension for the listener, which will have the overall effect of facilitating communication. This assumption of "listener accommodation" contains the additional assumption that speakers maintain a representation of the common ground they share with their interlocutors (H. Clark, 1996:43). They believe this representation is shared to a certain degree by all discourse participants (p. 49).

These assumptions raise two questions. First, how detailed is the representation of the common ground? If Jane has a conversation with her mother, does her model include the history of their shared experiences? Does she construct active representations of every object in their physical surroundings? Clark and Marshall (1981) suggest that communication follows certain "co-presence heuristics", such that discourse participants construct detailed models of the interlocutors, based on both their shared experience (linguistic and nonlinguistic) and their knowledge of each others' community membership. These proposed models are fairly detailed:

When Ann talks to Bob, she creates in memory a model of what is in Bob's mind - his knowledge, his perceptions, his current thoughts -- and she constantly updates it. Bob carries along a similar model of what is in Ann's mind. These models must include the right diary entries and encyclopedia chapters. Ann's model of Bob would contain all those chapters of her encyclopedia that correspond to communities she knows he belongs to.

(Clark and Marshall,1981:56)

APPENDICES

Clark and Marshall's argumentation supports the idea that speakers are able to access extremely specific knowledge about their interlocutors during conversation. On the other hand, memory limitations would prevent people from maintaining detailed models of their interlocutors in working memory. Do speakers represent every detail of the listener's knowledge in working memory, or is some information stored in a semi-accessible region in long-term memory, like Kintch's (1993) "extended working memory"? How much of this information is used on-line during language production?

There is evidence that speakers do not maintain models of every piece of information that is relevant for efficient communication. Anderson et al. showed, for example, that speakers speak less clearly during face-to-face conversations, when it was possible for the listeners to use visual articulatory information from faces (mouth gestures) to aid comprehension. However, they also showed that pronunciation did **not** vary according to whether speakers and addressees were looking at each other at a particular moment in the conversation. They concluded that speakers kept track of listener needs at a very gross level, but were not able to make finer, moment-to-moment distinctions. Further research on speakers' representations of common ground is important to understanding the relationship between reference comprehension and the choice of reference form in production.

A second question is whether Jane's choice of referring form is entirely driven out of consideration for her mother. Some production decisions may occur for "selfish" reasons, driven by constraints on planning and production. For example, in Arnold et al. (1998) we followed Wasow (1997) in arguing that constituent order is influenced at least in part by ease of production. With respect to patterns of reference and choice of referring form, it is important to discover whether any speaker choices are made with complete disregard for the listener's needs. The answer to these question would indicate the degree to which discourse patterns are driven by conscious, communicative strategies of the speaker, and the degree to which listener needs are taken into account in choice of reference form.

One area to study this issue is where the referring form is ambiguous with respect to the discourse context. With the factors discussed here, there is no a priori way to distinguish between the speaker's needs and the listener's. Since all five constraints (Recency, Subjecthood, Focus, Parallelism, and Goal Status) are properties of the linguistic context, an alternate story is possible: speakers could use pronouns only for the referents that are most activated in their own minds. Under this view, the correlation of each factor with Reference Form and Reference Continuation would have nothing to do with the listener's needs, but would merely reflect the fact that the things speakers talk about are active, so they use pronouns for them.

However, this story would not account for the tendency for speakers to use pronouns more often in situations where they are unambiguous. One feature of pronouns is that they provide very little information about their referents, so in many cases they may be ambiguous, as in 12a. However, often there is only one matching referent in the discourse situation, in which case there is no real ambiguity, as in 12b.

- (12) a. Nancy invited Alana to a party. She...
b. Nancy invited Ted to a party. She...

Francik (1985) investigated reference form by having people look at comics and describe the story. Some people saw stories about a boy and a man (same-gender condition), some saw stories about a girl and a man (different-gender condition). Overall, people used pronouns more often in the different-gender condition.

Francik's data would be difficult to account for without assuming that speakers are at least somewhat sensitive to the needs of their addressees. Assume that in both 12a and 12b, the speaker wanted to talk about Nancy in the second utterance. 12a and 12b are identical except for the gender of the second character. Thus, from the speaker's perspective, the referent for "Nancy" should be equally active in both cases. The only difference is the degree to which the referring form would be ambiguous. From the

APPENDICES

listener's perspective, on the other hand, the importance of ambiguity is obvious. The gendered pronoun provides unambiguous information about the referent in 12b, but not in 12a. Therefore, if the speaker's decision is influenced by the potential ambiguity of the referring form, it signals an attempt to facilitate comprehension for the listener.

Finally, the data and model I presented raise another question: do these factors influence on-line processes? I have argued that they are relevant because they facilitate or inhibit the process of anaphor resolution in language comprehension. However, with the exception of the experiment in §4.3., most of my data came from off-line measures. Other researchers have shown with on-line studies that anaphor resolution is influenced by Subjecthood (Gordon et al., 1993; Hudson-D'Zmura and Tanenhaus, 1998), Focus (Almor, 1995), Parallelism (Smyth and Chambers, 1996), Recency (Clark and Sengul, 1979), and Goal Status (Stevenson, 1995). The next step would be to show that these factors also influence on-line processes in the speaker's choice of referring forms in production.

5.4. Final words

The most important contribution of the present studies is that they reveal a commonality among Recency, Subjecthood, Focus, Parallelism, and Goal Status. Until now, these factors merely looked like a laundry list of things that affected anaphor resolution and choice of reference form. I have shown that they are all associated with an increased probability of subsequent reference to their referents. I suggested that this association leads them all to facilitate anaphor resolution. The reasons these patterns exist are somewhat different for each factor, but they all derive from generalizations about how people use language. My proposal predicts that choices in reference form should be influenced by any factor in a given language that is related to regular patterns of reference in that language. In sum, the cognitive processes involved in communication are the result of the social experience of individuals, and their ability to convert experience into knowledge for future use.