Evidence that Binding Principles Participate in a Constraint Satisfaction Process

William Badecker   &   Kathleen Straub
Johns Hopkins University   *University of Rochester*

Introduction

There is a recurrent view that language processing takes place in a cascade, with grammatical principles acting as an automatic filter on what information is made available to semantic and pragmatic processing mechanisms. Alternatively, one may view grammatical principles and semantic and pragmatic knowledge as applying more or less simultaneously (as the relevant information is encountered or accessed) to constrain the interpretation of an utterance. The domain of pronoun interpretation is a potentially fertile ground for addressing and evaluating these conflicting views both because there are well known grammatical constraints on pronoun interpretation and because there are several experimental techniques which allow one to examine various aspects of the interpretive processes.

Although the process of pronoun interpretation has received considerable attention in psycholinguistic research, few studies have focused on the role that grammatical principles play in the course of interpretation. A notable exception to this is the cross-modal priming study of pronoun and reflexives presented in Nicol and Swinney (1989). In that study, subjects listened to sentences like those in (1), and were asked to make lexical decisions about target words at a point simultaneous to the pronoun offset (indicated by a pound sign (#) in the example).

(1)   The boxer$_i$ told the swimmer$_j$ that the doctor$_k$ for the team would blame him$_{i/j/*k\ \#}$ for the recent injury.

Target words were either related to one of the subscripted nouns in (1) or were unrelated controls. For sentences like (1), Nicol and Swinney found priming for the targets related to the possible antecedents of the pronoun (boxer and swimmer) but not for the targets related to doctor, which would be ruled inaccessible to the pronoun by Principle B of the Binding Theory. Nicol and Swinney interpret this as evidence for initial grammatical filtering in that only accessible NPs are considered as potential antecedents by the pronoun interpretation mechanism[1]

---

[1] A schematization of this account for sentence (1) appears in Appendix A. Nicol and Swinney (1989) suggests that the primary results represent processing at some stage beyond the Binding Theory component in the diagram.

This poster presents experimental evidence that there can be processing effects for phrases that are excluded from the set of candidate co-referents by principles of Binding Theory. The first experiment demonstrates that the processing load associated with pronoun interpretation is greater for sentences containing two gender- and number-consistent candidate antecedents than for sentences in which only one such candidate is available. This ambiguity effect is exploited in the second experiment to show that the interpretation of a pronoun is not blinded to the content of a noun phrase which is inaccessible in virtue of Binding Theory principles.

Experiment 1

The first experiment was designed to assess effects of antecedent role and ambiguity on word-by-word reading performance. Previous studies using whole-sentence reading time measures and probe response measures have documented processing costs associated with ambiguous versus unambiguous pronoun contexts (Ehrlich, 1980; Garnham and Oakhill, 1985; Matthews and Chodorow, 1988), as well as a preference for subject over object antecedent (Grober, Beardsley and Caramazza, 1978; Crawley, 1986; Matthews and Chodorow 1988). This experiment attempts to replicate those earlier findings in order to validate the sensitivity of the self-paced, word-by-word reading paradigm to the processes of pronoun interpretation. Single-sentence test items were constructed in which the main clause subject and direct object were varied in grammatical gender to derive four pronoun interpretation conditions: no antecedent, subject antecedent, object antecedent, ambiguous (subject and object) antecedent. Examples of these conditions are shown in (2):

(2) no ant.:   Julie assured Lucy that he was prepared for the new job.
     sub. ant.:   Kenny assured Lucy that he was prepared for the new job.
     obj. ant.:   Julie assured Adam that he was prepared for the new job.
     ambig.:   Kenny assured Adam that he was prepared for the new job.

It was expected that if the word-by-word reading paradigm were sensitive to pronoun processing, one would observe longer reading times following the pronoun in the object antecedent and ambiguous antecedent conditions in comparison to the subject antecedent condition. Further, it is expected that this experiment will replicate the processing load effect for unheralded pronouns reported in Badecker and Straub (1992).

Methods
Subjects: Twenty-four undergraduates enrolled at the University of Rochester participated in this experiment to satisfy a class requirement.

Stimuli: Twenty-four sentence frames were constructed on the model of the items in (2) for the four pronoun-antecedent conditions to be contrasted: no antecedent, subject antecedent, object antecedent, and ambiguous antecedent. Prior to the self-paced reading presentation, a pencil-and-paper norming task was run in which subjects were presented individual test items and asked to indicate their preference for the gender of the antecedent of the embedded clause pronoun. In this way the possible biasing effects of the context leading up to the pronoun was assessed. Items were then classified as Subject Biased, Object Biased or Neutral Contexts for post-hoc analyses of reading time performance. Furthermore, in order to ensure that pronoun interpretation results were not disproportionately affected by the gender ambiguity of the potential antecedent NPs, main clause subject and object positions were filled with proper names that had been normed for gender recognizability. Male and female names were equated for length in letters, syllables and mean gender identification time and accuracy. Pronoun gender varied across items but remained constant within items across antecedent conditions. The four antecedent conditions were counterbalanced by item across four presentation lists. The 24 test sentences that each subject saw were pseudo-randomly distributed through a list of 116 sentences of varying construction type and length.

Procedure: Subjects were presented sentence items in a self-paced, word-by-word reading format. Words appeared in the center of the screen and were immediately replaced by the following word when the subject indicated that he/she was ready by pressing a predetermined button on a key-pad (right-hand response). Each sentence was followed by a probe word that differed from the sentence material in both text location and color. Subjects were required to indicate whether or not the probe word had appeared in the previous text passage by pressing designated keys for yes or no. This probe task was followed on intermittent trials by questions that were designed to encourage subjects to attend to the content of the test passages. Sentences were presented using the MEL text presentation routines. Reading times for each item in each sentence were collected. Additionally, accuracy for the probe items were recorded for later analysis.

Results

Reading time analyses were carried out for the four lexical positions leading up to and including the pronoun and for the four lexical items immediately following the pronoun. Only items in which the subject correctly responded to the probe item were included in the analysis of reading times. Reading times did not differ across the antecedent conditions for the lexical material leading up to and including the pronoun (for either average reading time across the four positions or for the individual word positions). However, as indicated in Figure 1, there were mean reading-time differences among the antecedent conditions for the positions following the pronoun.

The overall effect of antecedent condition was found to be significant by analysis of variance on subject analysis (F1(3,23)=5.29, p<.002) and approaching significance on item analysis (F2 (3,23)=2.32, p<.075). Further analyses were carried out contrasting reading times in the subject antecedent condition with object antecedent and ambiguous antecedent conditions.

Figure 1
Experiment 1: Mean reading time for post-pronominal words (positions 5-8) by antecedent position
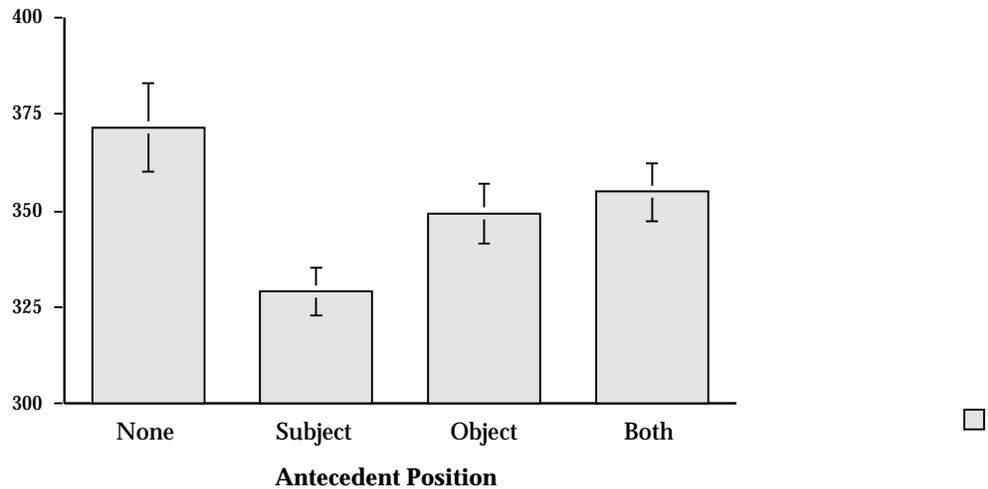
Analysis of variance revealed significant differences between mean post-pronominal reading times in subject and object antecedent conditions on both subject and item analyses ($F_1(1,23)=4.38$, p<.04; $F_2(1,22)=6.88$, p<.009, respectively). An examination of the reading times for Subject and Object antecedent conditions reveals that these differences are not attributable to the influence of contextual bias on the pronoun interpretation since, as indicated in Table 1, these differences are observed across the three Bias Categories.

Table 1: Reading times for post pronominal positions for Subject and Object antecedent conditions by context's antecedent bias.

	Subject Antecedent	Object Antecedent	
Subject Bias	310.9	325.0	-14.1
Neutral Bias	332.5	363.6	-31.1
Object Bias	340.1	348.4	-08.3

Although these contrasts do not rule out some contribution of context bias to the reading time measure, there is evidence of a subject antecedent preference across all three bias conditions.

Significant differences were also observed between Subject antecedent and Ambiguous antecedent conditions. Mean post-pronominal reading times differed on both subject and item analyses ($F_1(1,23)=4.90$, $p<.03$; $F_2(1,23)=3.94$, $p<.05$, respectively).

Discussion

Reading time comparisons in Experiment One showed that the self-paced, word-by-word reading paradigm is sensitive to differences in processing complexity associated with the interpretation of a pronoun within a sentence context. In particular, the study reveals reliable effects on pronoun construal for the number and position of gender-consistent NPs in a target sentence on pronoun construal. This ambiguity effect will be exploited in the second experiment to test the predictions of the Binding-Theory-as-Initial-Filter Hypothesis.

Experiment 2

On the expectations of the Binding-Theory-as-Initial-Filter Hypothesis the content of a phrase which is rendered inaccessible to a pronoun by principles of Binding Theory should have no effect on the pronoun's interpretation. That is, if both John or Joan are inaccessible to the pronoun he in a sentence like (3), there should be no processing cost that distinguishes rejecting these NPs as potential antecedents: Both NPs be filtered from the set of possible antecedents before the content of the NP can be evaluated by semantic-, pragmatic- or discourse-information based mechanisms.

(3)   a. Bob thinks that John will give him a better cut of venison next year.
        a. Bob thinks that Joan will give him a better cut of venison next year.

In contrast, if the Binding Principles are brought to bear on the interpretation of pronouns in a single constraint satisfaction process along with semantic, pragmatic and discourse mechanisms, then the primary consideration in determining whether the gender of an inaccessible NP can influence pronoun interpretation is determining when the various types of information become available from the linguistic input. On this view, the specification for gender of a noun phrase (along with other aspects of the NP's content) can contribute to possible antecedent selection independently of the Binding Theory considerations. This experiment was designed to evaluate the contrasting predictions of the Binding-Theory-as-Initial-Filter Hypothesis and the hypothesis that Binding Theory participates in pronoun-antecedent resolution along with other sources of information in a single constraint satisfaction process.

Methods

<u>Subjects</u>: Twenty undergraduates enrolled at the University of Rochester participated in this experiment to satisfy a class requirement.

<u>Stimuli</u>: Twenty-four sentence frames were constructed on the model of the items in (3). The gender of the accessible main clause subject NP and the inaccessible embedded clause subject NP were varied orthogonally yielding four presentation conditions. The two no-antecedent conditions contained a gender inconsistent NP in the main clause subject position while the two one-antecedent conditions, as exemplified in (3), contained a gender consistent NP in the main clause subject position. Within these conditions, the gender of the embedded subject was varied to derive gender consistent and gender inconsistent items. For the purposes of analyses, there were two major contrasts: (1) No antecedent versus one antecedent conditions; and, within the one antecedent condition, (2) Gender consistent versus gender inconsistent inaccessible NP. As in Experiment One, the names that alternated in the main and embedded subject positions were matched on letter length, syllable length and gender recognizability. The four conditions were counterbalanced by item across four presentation lists. The 24 test sentences that each subject saw were pseudo-randomly distributed through a list of 116 sentences of varying construction type and length.

<u>Procedure:</u> *See Experiment One*.

Results

      Reading time analyses of the one antecedent and no-antecedent conditions were carried out for the four lexical positions leading up to and including the pronoun and for the two lexical items immediately following the pronoun. Only items in which the subject correctly responded to the probe item were included in the analysis of reading times. Reading times did not differ across the antecedent conditions for the lexical material leading up to and including the pronoun (for either average reading time across the four positions or for the individual word positions). However there was a significant effect for the one- versus no-antecedent contrast on both subject and item analyses when mean reading times for the two post-pronominal positions were compared ($F_1(1,18)=5.10$, $p<.03$; $F_2(1,23)=4.79$, $p<.03$, respectively).

      A second set of analyses were computed comparing reading times on items in which the accessible, matrix subject NP was consistent with the pronoun gender and the inaccessible, embedded subject NP was either consistent or inconsistent with the gender of the pronoun. These reading times are shown in Figure 2.

Figure 2

**Experiment 2: Reading time for word position by gender
of inaccessible antecedent**
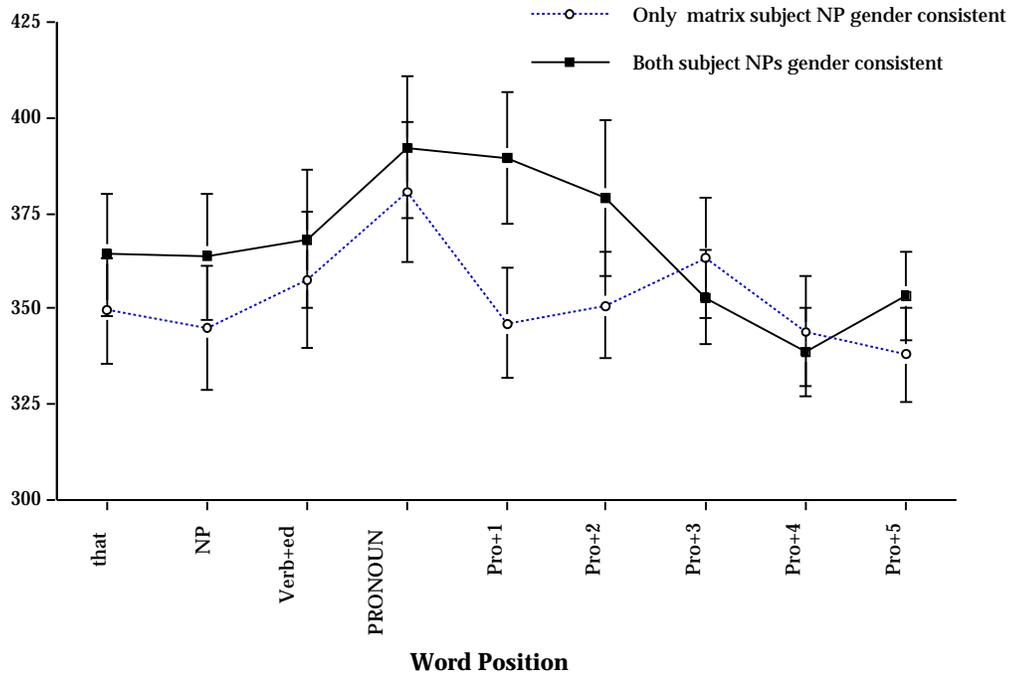
Reading times did not differ across the antecedent conditions for the lexical material leading up to and including the pronoun (for either average reading time across the four positions or for the individual word positions). However, mean reading times for the 2 post-pronominal positions on items in which only the matrix subject gender was consistent with the gender of the pronoun to be resolved were reliably faster on both subject and item analyses than those in which both the matrix and (inaccessible) embedded subject NP gender were consistent with the pronoun ($F_1(1,18)=13.01$, $p<.03$; $F_2(1,23)=4.42$, $p<.04$, respectively).

General Discussion

The results of Experiment Two provide a counter-evidence to the Binding-Theory-as-Initial-Filter Hypothesis. The effect for the content of the inaccessible NP indicates that the grammatical principles of Binding Theory do not act as an initial filter on the set of possible antecedents for pronouns. Instead, it appears that considerations of grammatical inaccessibility

enter into a single constraint satisfaction process that is initiated by the interpretive possibilities that are established by the referentially dependent lexical item (i.e., the pronoun). On this alternative view, the initial set of potential pronoun antecedents may be established by a variety of factors (including grammatical, semantic, pragmatic and discourse factors) which may or may not be in conflict with one another in antecedent selection. Just as local grammatical expectations do not constrain the set of meanings initially accessed for an ambiguous lexical item (Swinney, 1978; Seidenberg, Tanenhaus, Leihman and Bienkowski, 1981), grammatical expectations need not limit the initial set of candidate antecedents for a pronoun.

It should be noted that although our findings are incompatible with the interpretation that Nicol and Swinney (1989) offer for the priming patterns observed in their experiment, the results from these two studies are not themselves in conflict. Nicol and Swinney (1989) conclude that, since there is facilitation for the probe items related to both of the accessible candidate antecedents (boxer and swimmer in example (1), above) but not the probes related to the inaccessible NP (doctor in the example), the priming pattern reflects the set of candidates that the language processor considers--i.e., "all and only those referents that bear the appropriate syntactic relation to the referentially dependent NP." However, there are two alternative interpretations of the priming pattern they observe that are fully consistent with their results and the results reported here. The first, less interesting interpretation, is that facilitation in their experiment derives from antecedent selection, and the priming observed for the two accessible NPs reflects a mixture of two interpretive outcomes (those in which boxer is selected as the antecedent of the pronoun and those in which swimmer is selected). On this interpretation the observed difference between related item and control lexical decision target represents the mean of the facilitation induced by selecting the related item as the antecedent and the baseline performance obtained on trials in which the related item is not selected as the antecedent. The second interpretation is that the selective pattern of activation reflects a cumulative, though not final, product of the various of information types that contribute to antecedent selection. While some factors may contribute by adding to the activation level of the representation of a candidate antecedent, others may inhibit the activation level of one or more of the representations (Gernsbacher, 1989; McDonald and MacWhinney, 1990). In the event that a particular candidate antecedent receives positive activation from one information source and negative activation from another, one might expect that there would be an increase in the processing load (relating to the resolution of the pronoun) over what is observed when there is agreement among the various information sources engaged in antecedent selection. These expectations are born out by the findings reported here: When the content of an inaccessible NP is consistent with the gender of an unresolved pronoun, there is an observable increase in the processing load as

reflected by increased reading times for words immediately following the pronoun[2]. Further experiments will be carried out in order to examine which other factors may be observed to influence candidate antecedent evaluation for NPs in grammatically inaccessible positions (e.g., discourse focus).
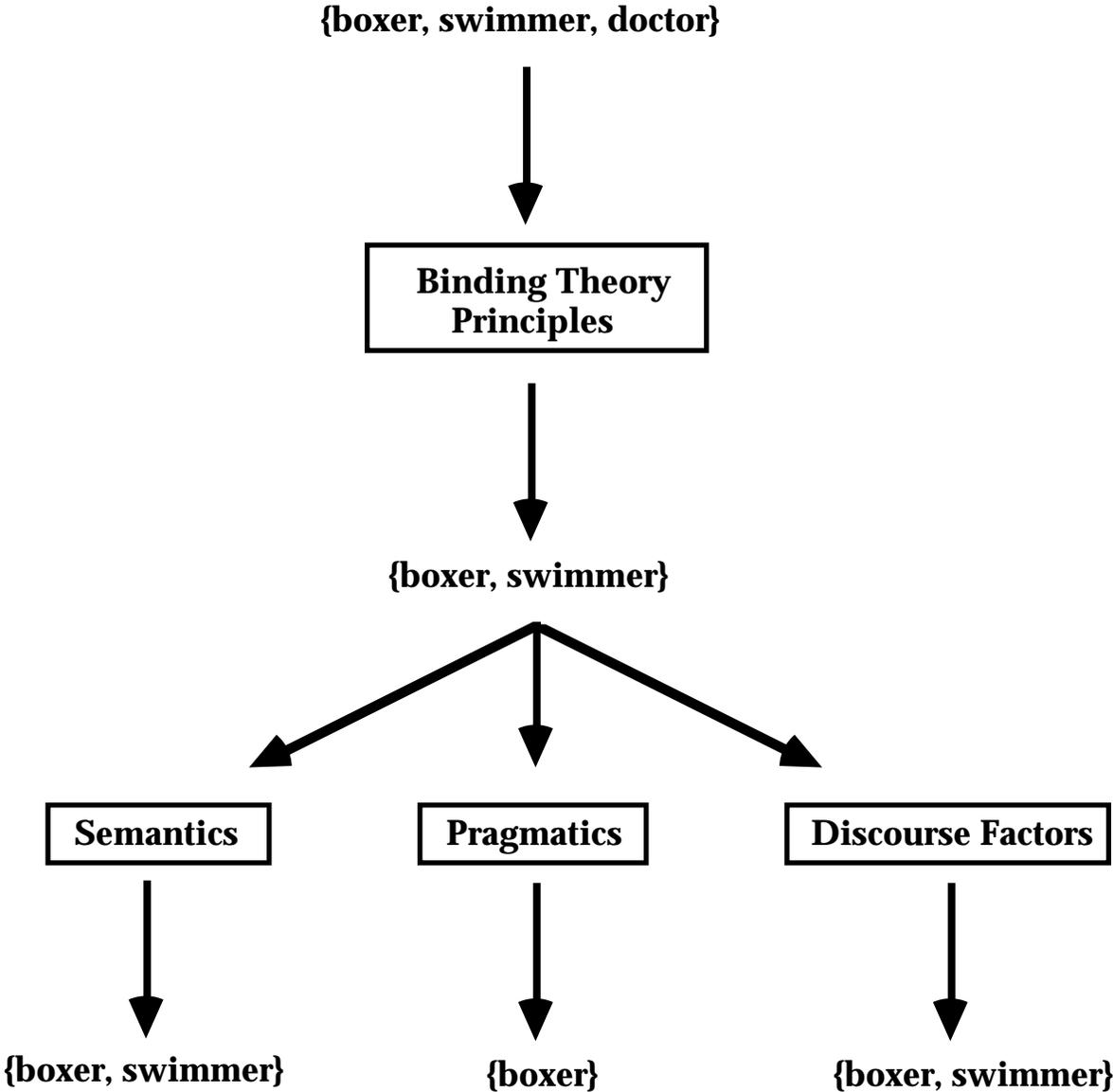
---

[2] This model is roughly schematized in Appendix B. A star (*) indicates a negative priming or suppression effect which might cancel out any postitive activation from semantic, pragmatic or discourse processing mechanisms. Appendix C depicts the Binding-as-a-Weighted-Constraint Model output for the for the minimal sentence pair examined in Experiment 2.   Figure C.1 represents the output of the model given the sentence input: Bob thinks that Joan will give him a better cut of venison next year. Figure C.2 represents the output using Bob thinks that John will give him a better cut of venison next year.   Note that whereas there is some positive activation for the inaccessible element in Figure C.2 as a result of processing based on various possible information types, only one participant receives positive activation in the processing situation depicted in Figure C.1. Compare these with the output of the Binding-Theory-as-an-Initial-Filter model using the same sentence inputs (shown in Appendix D, Figures D.1 and D.2, respectively).
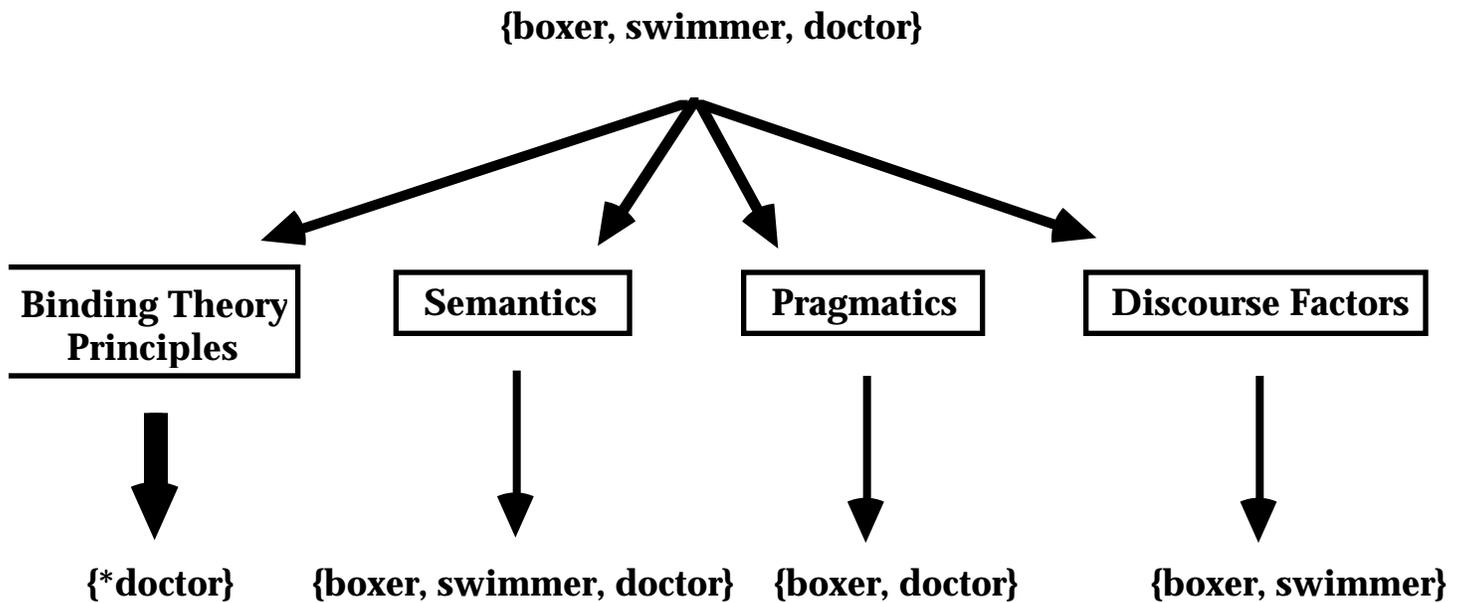
References

Badecker, W. and K. Straub (1992). Resolving pronoun-antecedent relations. Fifth CUNY Sentence Processing Conference, New York (March 19, 1992).

Crawley, R. (1986). Some factors influencing the comprehension of pronouns in texts. In C. Clifton (Ed.)   Proceedings of the Eighth Annual Conference of the Cognitive Science Society. Hillsdale, NJ, Erlbaum.   613-620.

Crawley, R. A. and Stevenson, R. J. (1990). Reference in single sentences and in texts. *Journal of Psycholinguistic Research* , *19*, 191-210.

Erlich, K. (1980). Comprehension of pronouns. *Quarterly Journal of Experimental Psychology* , 19, 245-264.

Garnham, A. and Oakhill, J. (1985). On-line resolution of anaphoric pronouns: Effects of inference making and   verb semantics. *British Journal of Psychology* , *76*, 385-393.

Gernsbacher, M. (1989). Mechanisms that improve referential access, Cognition, 32, 99-156.

Grober, E., W. Beardsley, and Caramazza, A. (1978). Parallel function strategy in pronoun assignment. *Cognition* , *6*, 117-133.

McDonald, M. and MacWhinney, B. (1990). Measuring inhibition and facilitation in pronouns. Journal of   *Memory and Language*, *29*, 469-492.

Matthews, A. and Chodorow, M. S. (1988). Pronoun resolution in two-clause sentences: Effects of ambiguity, antecedent location, and depth of embedding. *Journal of Memory and Language*, 27,   245-260.

Nicol, J. and D. Swinney (1989). The role of structure in coreference assignment during sentence processing. *Journal of Psycholinguistic Research 18*, 5-19.

Seidenberg, M., Tanenhaus, M., Leihman, J., and Bienkowski, M. (1981). Automatic access of the meanings of ambiguous words in context: Some limitations of knowledge-based processing. *Cognitive Psychology   14*, 489-537.

Swinney, D. (1979). Lexical access during sentence comprehension: (Re)Consideration of context   effects.*Journal of Verbal Learning and Verbal Behavior 18*, 645-659.

Appendix A: The Binding Theory as Initial Filter Model

{boxer, swimmer, doctor}

↓

| Binding Theory |
| Principles |

↓

{boxer, swimmer}

| Semantics | | Pragmatics | | Discourse Factors |

↓   ↓   ↓

{boxer, swimmer}   {boxer}   {boxer, swimmer}

The boxer told the swimmer that the doctor for the team
would blame him for the recent injury.

Appendix B: The Binding Theory as a Weighted Constraint Model

{boxer, swimmer, doctor}

Binding Theory Principles	Semantics	Pragmatics	Discourse Factors
{*doctor}	{boxer, swimmer, doctor}	{boxer, doctor}	{boxer, swimmer}

*The boxer told the swimmer that the doctor for the team
would blame him for the recent injury.*

Appendix C: The Binding Theory as a Weighted Constraint Model*

Figure C.1
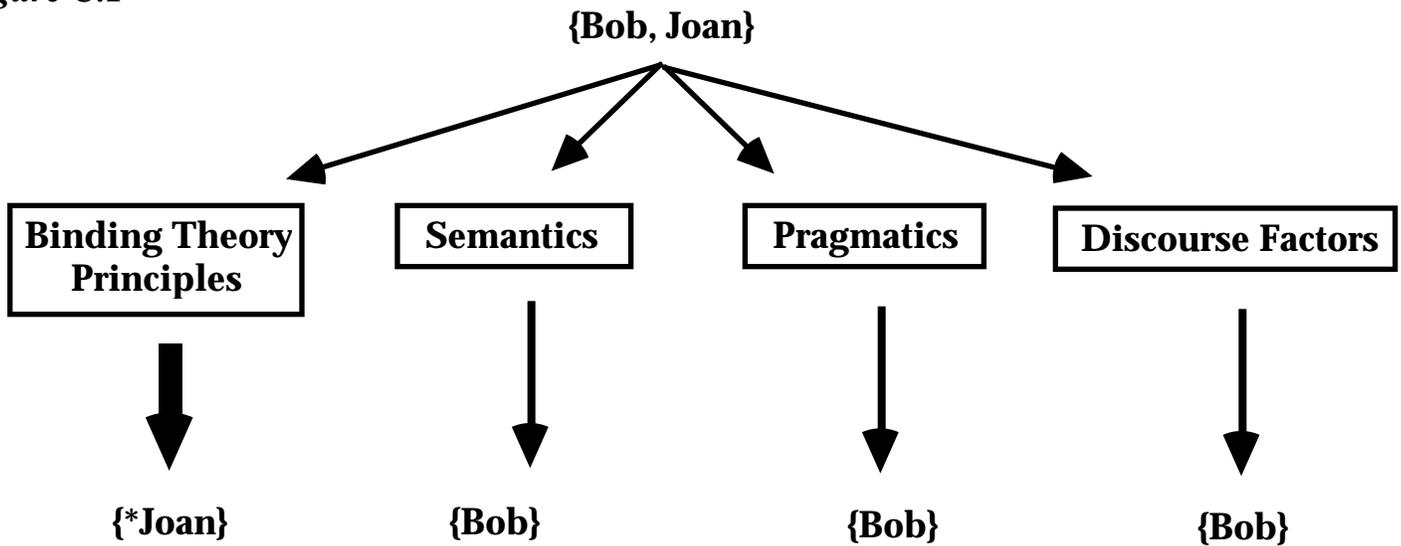
{Bob, Joan}

Binding Theory Principles	Semantics	Pragmatics	Discourse Factors
{*Joan}	{Bob}	{Bob}	{Bob}

Bob thinks that Joan will give him a better cut of venison next year.

Figure C.2

{Bob, John}

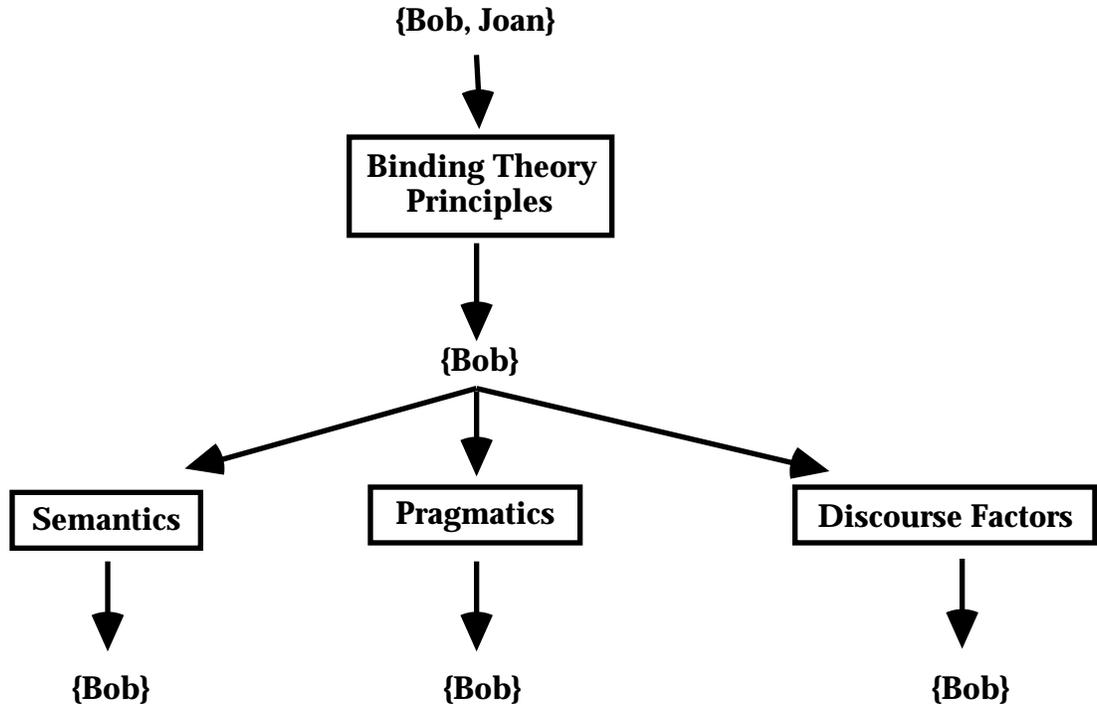| Binding Theory Principles	Semantics	Pragmatics	Discourse Factors
{*John}	{Bob, John}	{Bob, John}	{Bob}
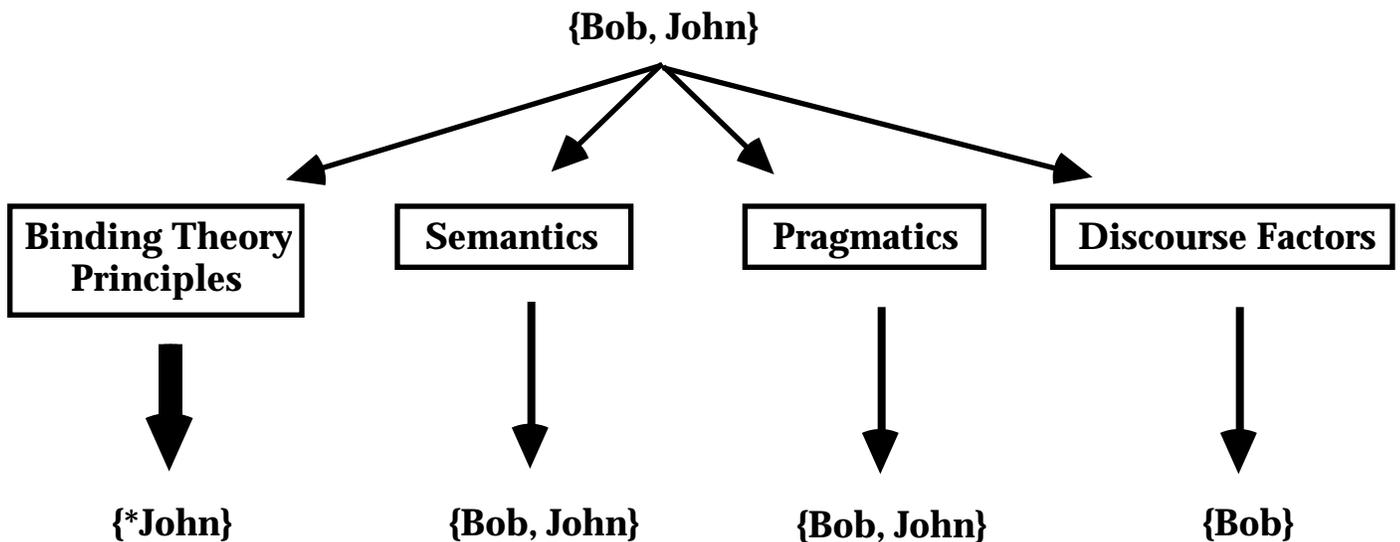
Bob thinks that John will give him a better cut of venison next year.

Appendix D: The Binding Theory as Initial Filter Model*

Figure D.1

{Bob, Joan}

Binding Theory Principles

{Bob}

| Semantics | Pragmatics | Discourse Factors |

{Bob}   {Bob}   {Bob}

Bob thinks that Joan will give him a better cut of venison next year.

{Bob, John}

Figure D.2

Binding Theory Principles

{Bob}

| Semantics | Pragmatics | Discourse Factors |

{Bob}   {Bob}   {Bob}

Bob thinks that John will give him a better cut of venison next year.