A Methodology for Extending Focusing Frameworks

Linda Z. Suri*
Fujitsu Nexion, Inc.

Kathleen F. McCoy†
University of Delaware

Jonathan D. DeCristofaro‡
University of Delaware

We address the problem of how to develop and assess algorithms for tracking local focus and for proposing referents of pronouns. Previous focusing research has not adequately addressed the processing of complex sentences. We discuss issues involved in processing complex sentences and review a methodology used by other researchers to develop their focusing frameworks. We identify difficulties with that methodology and difficulties with using a corpus analysis to extend focusing frameworks to handle complex sentences. We introduce a new methodology for extending focusing frameworks, which involves two steps. In the first step, a set of systematically constructed texts are used to identify an extension of the focusing framework to handle a particular kind of complex sentence. In the second step, a corpus analysis is used to confirm the extension. We explain how our methodology overcomes the difficulties faced by other approaches.

1. Introduction

The central problem addressed in this work is how to develop and assess algorithms for tracking local focus and for proposing referents of pronouns for use in natural language processing (NLP) systems.[1] By "local focus," we refer to the person, object, property, or concept that a sentence is most centrally about within the discourse context in which it occurs. The appropriate movement and marking of local focus, and the appropriate choice of the form of a noun phrase (NP) based on local focus information, are considered to contribute to the local coherence exhibited by discourse (Sidner [1979], Grosz, Joshi, and Weinstein [1983, 1995], Carter [1987], and others).

In addition, local focus information is one source of information that is used by readers and hearers for interpreting pronouns. Some researchers (e.g., Hobbs 1978; Lappin and Leass 1994) have proposed pronoun resolution algorithms that do not involve focus tracking. However, our view is that local focus tracking and pronoun resolution are mutually dependent processes. The local focus information influences pronoun resolution, and pronoun resolution, in turn, influences updating focus information. Therefore, the tracking of local focus is crucial for the interpretation of pronouns.

* 1807 Park 270 Drive, Suite 350 St. Louis, MO 63146. E-mail: suri@nexen.com
† Department of Computer and Information Sciences, University of Delaware, Newark DE 19716. E-mail: mccoy@cis.udel.edu
‡ Department of Computer and Information Sciences, University of Delaware, Newark DE 19716. E-mail: decristo@cis.udel.edu

There have been several algorithms described in the literature for tracking local focus information and for using this tracked information to do pronoun resolution. In this paper we first briefly introduce the notion of local focusing and what a local focusing algorithm is intended to capture. Generally the way that the focus of a sentence is expected to shift through a discourse is dependent on some syntactic properties of the sentence. However, most of the work on tracking local focus has concentrated on simple (single clause) sentences. Thus previous work on focusing has not adequately addressed the processing of complex (i.e., multiclausal) sentences. We discuss a number of issues involved in the processing of complex sentences in order to motivate the need for a methodology for extending focusing frameworks to handle them. We review a methodology used by other researchers to develop their focusing frameworks, and we identify some difficulties with that methodology. We examine the possibility of using a corpus analysis to extend a focusing framework, and briefly describe potential problems with such an approach. We then introduce our own two-part methodology for extending focusing frameworks, which we call the Semantically Slanted Discourse (SSD) Methodology. The first part of the methodology consists of an exploratory phase in which possible extensions to a focusing algorithm are discovered through the use of carefully constructed discourses that rely on the potential tension between focusing and world-knowledge factors in pronoun resolution. We show how this first phase can be used to propose an extension to a local focusing framework in order to handle a given type of complex sentence. The first phase is then followed by a corpus analysis to confirm its findings. We explain why a corpus analysis used to confirm an extension is more practical than one used to identify an extension in the first place.

2. What is Local Focusing?

We use the term "local focusing framework" to refer to a theory or framework consisting of a (set of) focus tracking algorithm(s) and a (set of) pronoun resolution algorithm(s). A local focusing framework records and makes use of information about **focusing factors** and indicates how these factors influence (intersentential) pronoun resolution. Generally, focusing factors include such things as:

1. Grammatical role. In several focusing frameworks (Sidner 1979; Grosz, Joshi, and Weinstein 1983, 1995), some grammatical roles (e.g., surface subject and surface direct object) are considered indicative of what is in focus. Also, a pronoun resolution algorithm might prefer to find an antecedent from a previous sentence that has the same grammatical role as a pronoun in the current sentence.

2. Pronoun use. Most, if not all, focusing frameworks are constructed on the basis of the belief that pronominalization is often indicative of focus. At the same time, when an intersentential pronoun is found, a pronoun resolution algorithm generally prefers to find an antecedent that is highly focused in the previous sentence.

3. Constancy of focus. Many focusing frameworks assume that if an item that was the focus of the previous sentence occurs in the current sentence, it is more likely to be focused in the current sentence. As above, a pronoun resolution algorithm that suggests the focus of the previous sentence as the antecedent of a pronoun in the current sentence would be consistent with this factor.

4. Shifting preferences. Most, if not all, focusing frameworks specify preferences for how focus is most likely to shift if the focus of the current sentence is not the same as the focus of the previous sentence.

5. Clue words and phrases. Most, if not all, focusing research takes into consideration that clue words and phrases might affect pronoun resolution and what is focused in a sentence.

6. Syntactic form. In some frameworks, certain syntactic forms (e.g., clefted sentences) might indicate an item as highly focused.

The specific list of focusing factors and the way they interact are different for different focusing frameworks. A particular focusing framework must identify a set of focusing factors and must indicate how these focusing factors interact to suggest cospecifications for pronoun resolution and to identify the focus of a sentence.

A local focusing framework is not intended to *independently* interpret pronouns. Rather, a local focusing framework is intended to suggest cospecifications for pronouns in a reasonable order. An inferencing mechanism that makes use of **semantic factors** (such as semantic case constraints, world knowledge, rhetorical relations, etc.) must be used to confirm or reject a suggested cospecification. Thus, local focusing frameworks are intended to capture a coherence factor in discourse that influences preferences for how to resolve pronouns independent of semantic factors.

3. Processing Complex Sentences: A Reason for Extending Focusing Algorithms

Although complex sentences are prevalent in written English, most other local focusing research (focusing: Sidner [1979] and Carter [1987]; centering: Grosz, Joshi, and Weinstein [1983, 1995], Brennan, Friedman, and Pollard [1987], Walker [1989, 1993], Kameyama [1986][2], Walker, Iida, and Cote [1994], Brennan [1998], Kameyama, Passonneau, and Poesio [1993], Linson [1993] and Hoffman [1998]; and PUNDIT: Dahl [1986], Palmer et al. [1986], and Dahl and Ball [1990]) did not explicitly and/or adequately address how to process complex sentences. Thus, there is a need to extend focusing algorithms.

An exception to this rule is the work of Strube (1996) (which applies functional-information-structure-based criteria on a per-clause basis), Kameyama (1998), and Strube (1998). Kameyama's focus was on intrasentential anaphora and she attempted to define an "utterance" in the face of syntactic complexity. Strube, in very recent work (Strube 1998), handles arbitrary sentence complexity. Still, neither considers how particular types of complexity might affect a broad range of focusing factors.

Notice that there are a number of ways that a given type of complex sentence might be handled by a focusing algorithm. For instance, consider processing a complex sentence of the form "SX because SY," where SX and SY each consist of a single clause. One might imagine processing the SX clause and then the SY clause (i.e., resolving the pronouns in these clauses and updating the focusing data structures) as if the clauses were a sequence of simple (i.e., single clause) sentences. On the other hand, it may be the case that for this type of complex sentence, the sentence should be treated as a single unit of processing with elements of one of the clauses dominating the

---

2 Kameyama (1986) did address the issue of multiple subjects, but did not address the general problem of developing a methodology for determining how a focusing framework should handle complex sentences.

processing. (For further discussion of these and other possible processing possibilities, see Suri [1993].)

The question we address is how one can appropriately extend a focusing mechanism to handle various kinds of complex sentences. We illustrate our approach by focusing on one kind of complexity: sentences of the form "SX because SY" where SX and SY are simple sentences. In Section 7 we discuss the application (or generalization) of this methodology to other types of complex sentences.[3]

4. A Methodology Used in Other Local Focusing Work

Recall that local focusing theories are attempting to capture patterns of focus movement and patterns of relations between anaphors and their antecedents that are independent of semantics, world knowledge, rhetorical relations, etc. Because of this, the method for determining how to process particular kinds of complex sentences that might seem the most natural is to construct semantically neutral discourses[4] that involve the type of complex sentence under study, and gather linguistic judgments to determine how people prefer to resolve the pronouns. In fact, in exploring *other aspects* of local focusing frameworks, other literature appears to have tried to make use of semantically neutral texts in this fashion (e.g., Brennan, Friedman, and Pollard 1987; Walker, Iida, and Cote 1994). However, in trying to construct discourses to determine how to process a particular kind of complex sentence, we realized it is difficult to construct discourses that are truly semantically neutral *and* sound natural. This task is further complicated by the need to construct a number of semantically neutral texts in order to control for and isolate each of the factors that might affect how readers prefer to resolve pronouns.[5] These factors include the influence of the other complex sentence structures in the discourse, and the factors that affect focus computation and pronoun resolution for simple sentences, i.e., focus history, the syntactic roles of pronouns and their potential antecedents, verb aspect and tense, and so on.

More importantly, when one is constructing texts *without the benefit of a systematic methodology*, one cannot be sure that the collection of constructed texts are representative of naturally occurring text in terms of the interactive relationships within and across focusing and semantic factors, and their influence on pronoun interpretation. As a result, there is a danger of tuning a theory to handle a discourse phenomena that is the exception rather than the norm in naturally occurring situations.

5. Using a Corpus Analysis to Extend Frameworks

Because of the problems associated with using constructed discourses, it is natural to turn to some kind of corpus analysis to extend a focusing framework. For example, one might measure how well an extension of a framework handles a type of complex sentence by measuring how accurately and efficiently it suggests referents for pronouns in texts that contain the type of complex sentence under consideration. One could count how often the extended framework suggests a wrong referent (which would not be rejected by an ideal inferencing mechanism), and how many referents it suggests (on average) to the inferencing component before the correct referent is selected.

---

3 Note, it is possible that various types of complex sentences would each need to be handled differently by the focusing algorithm.

4 I.e., discourses whose pronouns cannot be unambiguously resolved on the basis of semantic/world-knowledge factors alone.

5 This need was not addressed by previous focusing work in an adequate or systematic fashion.

Such a corpus analysis has been used to compare pronoun resolution algorithms that are not based on a focusing framework. For instance, Hobbs (1978) proposes a simple pronoun resolution algorithm that proposes referents for pronouns that it finds by walking a parse tree of a sentence (and the previous sentences) in a particular order. He examined several hundred pronoun occurrences in a variety of texts to show how well the algorithm was able to identify the correct referent.

Lappin and Leass (1994) developed a pronoun resolution algorithm that chose a referent for a pronoun from a set of potential referents (which were filtered to assure their syntactic and morphological appropriateness) on the basis of a "salience factor" rating. The rating for a potential referent was raised if, for example, it was a subject, it occurred recently (the algorithm, like Hobbs's, preferred intrasentential pronouns), it was a head noun, or it had a parallel grammatical role with the pronoun. This algorithm was also tested using a corpus analysis (and it was compared to Hobbs's algorithm with a corpus analysis as well).

Strube (1998) introduced a novel pronoun resolution algorithm that handled intra- and intersentential anaphora uniformly. He was able to apply his algorithm to a corpus and showed that his algorithm outperforms that of Brennan, Friedman, and Pollard (1987) using the definition of utterance given in Kameyama (1998).

Notice that none of the factors on which the above algorithms were based explicitly take into account the specific complexity exhibited by a sentence. (For example, there is no difference in the set of factors for an "SX because SY" sentence as compared to an "SX although SY" sentence, even though it is possible that the subsequent pronouns should be resolved differently in the two cases.) In addition, neither Hobbs (1978) nor Lappin and Leass (1994) tracked focusing information. While Strube (1998) is closer to a "focusing-based" algorithm, it considers very few of the focusing factors discussed in the previous section. On the other hand, pronoun resolution algorithms dependent on the kind of focusing information discussed in the previous section are affected by complex sentence structures. Thus we must face the challenge of determining how the focusing algorithm (and pronoun resolution algorithms) should act in the face of the complexity introduced by a type of complex sentence.

Other researchers have attempted to make use of corpus analyses in focusing research (though they did not attempt to extend the various algorithms to handle complex sentences). Walker (1989) performed a corpus analysis on written and spoken English to compare centering with Hobbs's algorithm (Hobbs 1976) in terms of their accuracy and coverage for finding the cospecifiers of pronouns. Walker also performed a corpus analysis (on spoken English) to investigate how centering should process a sentence beginning with the word *Now*, which she assumes (frequently) marks a new discourse segment (Walker 1993). Linson (1993) analyzed a corpus of spoken data to investigate focus transition patterns. (Previous work on centering assumed a particular priority on focus transition possibilities.) However, in each of these analyses, the researcher did not specify how complex sentences in the corpora were processed by the centering algorithm. Because the complex sentences may affect the way the centering data structures should be computed, it is not clear how the reported results should be interpreted.[6]

While we believe that corpus analyses are ultimately necessary for evaluating focusing frameworks, any analysis based solely on a corpus analysis will be faced with several potential difficulties.

---

6 In fact, in some cases, Linson treated a complex sentence as two units for processing, and in others, he treated a complex sentence as a single unit for processing (Suri 1993).

First, a focusing framework is intended to capture a reader's preferences for focus movement and pronoun resolution *independent* of world knowledge, semantics, and other pragmatic factors. Very large amounts of text would have to be analyzed to control the influence of these factors; yet, since there are few tools available for this type of analysis, this task would be formidable.

Second, a corpus analysis may be useful in *comparing* various extensions of a focusing framework, but it is up to the designer to decide *which* extensions to compare.[7] Using a corpus analysis alone, a novel extension cannot emerge by becoming evident as a side effect of the analysis. Generally speaking, before a corpus analysis can be used, the researcher must have made all decisions concerning the processing of all types of sentences occurring in the corpus. These decisions would include:

1. Whether and how a complex sentence should be segmented for processing (for pronoun resolution and focus computation). (Note that the answer to this might vary for different types of complex sentences.)

2. Whether at a given point in the discourse, the framework should prefer for the focus to remain the same or for the focus to shift in some particular manner.

3. What and how linguistic information about the text should affect the (predictive) pronoun resolution algorithms and updating focusing data structures.
   One *might* include the following items among this information:

   - grammatical roles of elements
   - syntactic form of clauses
   - the syntactic from of the sentence as a whole
   - pronominalization (in the current sentence)
   - clue words
   - verb aspect and tense

   One must also determine how these factors should be taken into account in the algorithms.

4. How factors discussed in items 2 and 3 interact in pronoun resolution and in updating focusing information.

The decisions about the matters discussed in the above list would constitute a (version of a) focusing framework. Thus, when one tries to compare different extended versions of a focusing framework one is effectively asking how the different sets of possible decisions about the matters discussed above compare to one another.

To perform a corpus analysis to identify the appropriate extension of a framework in order to process a particular type of complex sentence, all possible extensions of the framework (which are to be tested in the corpus analysis) must be completely identified *prior* to the corpus analysis. Notice, however, that there is no guarantee that the *correct* extension will be specified and tested in the corpus analysis approach. One might overlook the appropriate answers to how to segment sentences and how to process a particular kind of complex sentence. In addition, the number of possible

---

7 To use a corpus analysis and learning algorithm to automatically learn an extension, not only must the corpus be marked with the referents of all the anaphors, but also all the focusing data structures must be specified for each sentence. This task faces many of the problems outlined below.

extensions is likely quite large, and thus the number of alternative corpus analyses will likely be quite prohibitive.

Perhaps the most significant and problematic obstacle for determining how to extend a focusing framework to handle a particular kind of complex sentence via a corpus analysis is the following: if one does not know how to process many types of complex sentences, it is difficult to perform a corpus analysis to determine how to process a given type of complex sentence, since instances of that type of complex sentence are likely to be preceded and followed by other types of complex sentences. In particular, if we are trying to determine how to process a particular type of complex sentence, and that type of sentence occurs as sentence S_n in a text in a corpus, we do not want misconceptions about how to process sentence S_{n-1} or how to process sentence S_{n+1} in the text to mislead us in deciding how to process a complex sentence like sentence S_n. Furthermore, many sentences in the corpus are likely to involve multiple levels of complexity. Thus, it is very difficult to isolate the influence of one complex sentence structure from the influence of another complex sentence structure when performing a corpus analysis.

In sum, in order to perform a corpus analysis, it is necessary to make many decisions prior to the analysis. On what basis should these decisions be made? There are many factors that affect focusing and pronoun resolution that need to be studied and yet there is no way to systematically isolate these factors from each other in naturally occurring text. Also, one might not easily find a portion of text with an appropriate combination of features that is needed to test one's hypothesis. Furthermore, the testing of a hypothesis is likely to require looking at many portions of text with different combinations of values for focusing factors (e.g., pronominalized subject versus not pronominalized subject, whether a subject is coreferring with the subject [or some particular focusing data structure] of the preceding sentence). In addition, in order to perform a corpus analysis to determine how to extend a focusing framework, one must make decisions about how to segment and process all kinds of complex sentences that occur in the corpus. This requires knowing a priori what all of the processing possibilities are, or what all of the possible extensions to the framework are, and it requires testing all of these possible extensions.

We feel that these difficult problems make it infeasible to determine how to process complex sentences using a methodology based on a corpus analysis alone. On the other hand, once a possible extension is identified, verification with a corpus analysis is a necessary step to be sure that a found extension is appropriate. We discuss how to perform such a corpus analysis and indicate some specific problems with using a corpus analysis in the context of this problem in Section 6.5.

6. Our Two-Part Methodology for Determining How to Process Complex Sentences

As we have pointed out, there are several potential problems with analyses using constructed discourses and with corpus analyses. Our methodology combines specific instances of both of these methodologies; these specific instances were designed to overcome the difficulties we identified.

The first part of our methodology involves systematically constructing discourses (of a type to be described) and gathering acceptability judgments on these discourses. The discourses are constructed in such a way as to help identify a plausible extension of a focusing algorithm that would handle the type of complex sentence in question. The resulting extension must then be confirmed by a corpus analysis to ensure that the constructed discourses properly represent all influences actually found in naturally occurring text.

An alternative to the corpus analysis phase is to perform psycholinguistic experiments, such as those of Gordon, Grosz, and Gilliom (1993) and Hudson-D'Zmura and Tanenhaus (1998), which validate aspects of centering theory by measuring subjects' reading times of several types of sentences. This is a reasonable approach since, in those studies, the theory being validated, centering, was already given. In our case, the second phase is trying to verify a hypothesis suggested by the human subjects' judgments (collected in less-controlled circumstances) from the first phase. So, either a new set of test discourses or a new set of human subjects would be needed in order to execute a psycholinguistic experiment. For the sake of simplicity and to avoid the strict constraints on using human subjects, it is more practical for us to rely on a corpus analysis in the verification phase. Furthermore, the corpus analysis reveals how language is actually used in practice, rather than depending on a small set of discourses presented to the human subjects.

6.1 Semantically Slanted Discourse (SSD) Methodology: The Motivation for the First Part

In previous literature on local focusing (e.g., Sidner 1979; Grosz, Joshi, and Weinstein 1983, 1995; Brennan, Friedman, and Pollard 1987), researchers used a small number of constructed texts to justify aspects of their focusing frameworks and to assess and compare focusing frameworks. However, they did not explicitly address how one should construct sets of texts in order to draw accurate conclusions about local focusing. The first part of our methodology is intended to help the researcher construct sets of texts (i.e., minimal pairs or minimal quadruples) that allow components of a focusing framework to be systematically isolated and thus allow one to appropriately assess focusing frameworks.

To appreciate the reasoning behind the first part of our methodology, or what we call our Semantically Slanted Discourse (SSD) Methodology, recall that local focusing frameworks are intended to capture the preferences for pronoun resolution *independent of* semantics, world knowledge, rhetorical relations, and other kinds of pragmatics. Thus, they are intended to capture how one would resolve pronouns and update focusing information in discourse that is neutral in terms of these factors. Presumably in such texts, *only* focusing factors would affect pronoun resolution. In a semantically nonneutral discourse, semantic factors (semantics, world knowledge, etc.) can override the preferences of the focusing framework by rejecting potential referents proposed by a focusing framework.

Taking this into account, in order to determine how best to process a particular type of complex sentence, we decided to construct discourses that have two important properties:

1. The set of discourses must be systematically constructed to ensure that each possible combination of focusing factors is represented in a discourse. Below, we show how this is done for a particular focusing framework and a single type of complex sentence.

2. The discourses are intentionally loaded or slanted for pronoun interpretation based on semantic factors. We call such a discourse "semantically slanted" because the interpretation of all of the pronouns is fully determined by the semantic factors alone.

We contend that in a semantically slanted discourse, if the text seems ambiguous or awkward, or if one needs to reinterpret a pronoun, then the focusing preferences

for pronoun resolution are at odds with the preferences based on semantics, world knowledge, or other pragmatic factors. On the other hand, if the text seems acceptable/natural, then we contend that the preferences for pronoun resolution based on focusing agree with preferences based on semantic slanting. Thus, gathering acceptability judgments about these systematically constructed semantically slanted discourses should help us identify what the focusing preferences are, and thus how a focusing framework should be extended to handle a given type of complex sentence. This is the idea at the heart of our methodology.

6.2 Isolating the Complexity

In using semantically slanted discourses to uncover an extension of a focusing algorithm, discourses must be constructed which:

1.   isolate the complexity under study

2.   (when taken together) determine an extension of the focusing algorithm by determining how the focusing factors, the aspects of the text taken into consideration by the focusing algorithm, should "behave" in the face of the complexity under study

The first of these issues influences the overall form of the discourses being constructed. The second requires systematically constructing a number of discourses that alter the various focusing factors in such a way as to isolate their potential influence.

In order to isolate the complexity under study, we construct a set of discourses of the following form, for which the interpretation of the NPs is fully determined by the semantics of the text and world knowledge:

Example 1
(S1)   Simple sentence

(S2)   Sentence with one level of complexity (i.e., having two clauses), introduced by the syntactic form of interest.

(S3)   Simple sentence

In examining linguistic judgments about such texts, our goal is to identify preferences imposed by the syntactic form of S2 for:

•   resolving pronouns in S2

•   updating the focusing data structures after S2 so that the pronouns of S3 can be correctly resolved in a manner that is consistent with resolving pronouns in a sentence following a simple sentence or another kind of complex sentence.

The motivation for having S1 be a simple sentence is to avoid any effect a complex sentence might have on the focusing data structures going into S2. Similarly, the motivation for having S3 be a simple sentence is to avoid any effect that a complex sentence structure in S3 might have on pronoun resolution in S3.

6.3 Systematic Construction of Discourses for "SX because SY" Sentences

Our methodology calls for the systematic construction of a set of discourses that ensures that all possible combinations of the focusing factors are represented. This will allow an extension of a given focusing framework to emerge (since the set of discourses essentially captures all possibilities for an extension). To better illustrate the

types of discourses our methodology calls for constructing, let us consider what would be needed to extend a particular focusing framework, RAFT/RAPR (described in Suri [1993]), to handle resolving subject pronouns in sentences of the form "SX because SY" where SX and SY are simple sentences, and in a sentence following that type of sentence.[8]

But, first, let us briefly introduce some facts about the RAFT/RAPR algorithm. RAFT/RAPR defines a data structure called the Subject Focus, abbreviated SF, which, for a simple sentence, is taken to be the subject. The algorithm prefers to resolve a subject pronoun (in a simple sentence) so that it corefers with the contents of the Subject Focus of the previous sentence. When processing a simple sentence, if the suggested referent for a subject (i.e., the previous sentence's Subject Focus) is rejected by inferencing with world knowledge, semantics, and pragmatics, then other elements in the previous discourse are tried in a specified order; this ordering is (indirectly) influenced by such things as whether a pronoun was used in the previous sentence (since pronouns are indicative of focus and therefore influence the computation of focusing data structures for the previous sentence). Notice that the algorithm prefers constancy in Subject Focus over shifting the Subject Focus.[9]

To address the question of how to process "SX because SY" sentences, we examined discourses of the form:

Example 2
(S1)   simple sentence

(S2)   SX because SY

(S3)   simple sentence

The set of constructed discourses were "variations" of the form shown in Example 3.

Example 3
(S1)   Dodge was nearly robbed by an ex-convict the other night.

(S2)   [Dodge] captured [the ex-con] because [the ex-con] was so stupid and clumsy.

(S3)   Then [Dodge] called the police.

Notice in this example that SX = "[Dodge] captured [the ex-con]" and SY = "[the ex-con] was so stupid and clumsy". We constructed variations of this text to tease out how the various focusing factors interact when processing an "SX because SY" sentence or a simple sentence following an "SX because SY" sentence. The constructed variations of the text (described below) changed the form and referents of the text in square brackets "[ ]". In order to keep the semantic slanting appropriate, these changes caused the content of the text to be changed as well.

---

8 Throughout the paper, except in Section 6.5, we use "SX" and "SY" to denote simple sentences or simple clauses.
9 The centering algorithm prefers for the center to remain the same rather than to change, and it also prefers for the center to be the subject of the sentence. In general, focusing algorithms have a preference for focus/foci to remain the same rather than to change; this preference seems consistent with the goal for text to be coherent.

Some questions that must be answered in coming up with an extension of RAFT/RAPR for handling "SX because SY" sentences are:

1. If the Subject(SY) is a pronoun, how should Subject(SY) be resolved? Recall that RAFT/RAPR would prefer a subject pronoun to corefer with the subject of the previous sentence. In the face of this complexity, should the algorithm prefer that the Subject(SY) corefer with Subject(S1) or with Subject(SX)?

2. How should Subject(S3) be resolved? I.e.,

   • Preferring Subject(SX) always? (This would suggest that RAFT/RAPR should compute the Subject Focus of the "SX because SY" sentence to be Subject(SX).)
   • Preferring Subject(SY) always?
   • Preferring Subject(SX) or Subject(SY) depending on which is pronominalized?
   • Preferring Subject(SX) or Subject(SY) depending on which is coreferential with Subject(S1)?
   • Based on some other preference?

The answers to these (and all such similar questions) constitute a decision about how and whether the complex sentence should be segmented, and how to weigh the influences of the various focusing factors such as pronominalization and focus history in resolving pronouns and in updating the focusing data structures.

To see how we make up the discourse variations to cover all processing possibilities, consider this abstract view of the text, which indicates the NPs we are interested in:

Example 4
(S1)   Subject(S1) . . . Direct-Object(S1)

(S2)   Subject(SX) . . . Direct-Object(SX)
        because Subject(SY) . . .

(S3)   Subject(S3) . . .

To find an extension of the focusing algorithm, we need to construct text variations that capture all the different ways the grammatical roles, focus history, and pronominalization (i.e., the focusing factors) might interact in determining the referent of the pronoun in Subject(S3).[10] We construct text variations corresponding to variations of the following parameters:

1. Whether Subject(S1) is the ex-convict or Dodge. ("An ex-convict nearly robbed Dodge the other night" vs. "Dodge was nearly robbed by an ex-convict the other night.")[11]

---

10 It is crucial that we keep the *form* of the complex sentence constant (i.e., "SX because SY") but manipulate the text so that the focusing data structures before processing S2 take on all possible values. This allows us to identify how these data structures should influence pronoun resolution within S2 and S3.

11 Passive sentences are handled by most focusing algorithms; however, see Section 6.4.1 for a discussion of why the choice between active and passive voice could not account for the findings of the experiment.

- Note that the Direct-Object(S1) will always introduce the other actor. This helps test the focusing factor of grammatical role.

2. Whether Subject(SX) of S2 is the ex-convict or Dodge. ("[Dodge] captured [the ex-convict] because [the ex-convict] was so stupid and clumsy" vs. "[The ex-convict] woke [Dodge] up because [the ex-convict] was so stupid and clumsy.")

   - Notice that 1 and 2 together will vary whether or not the Subject(S1)=Subject(SX).
   - Note if Subject(S1)≠Subject(SX) then Direct-Object(S1)=Subject(SX), again testing grammatical role effects.

3. Whether Subject(SY) of S2 is the ex-convict or Dodge. ("[The ex-convict] tied [Dodge] up because [the ex-convict] didn't want any trouble" vs. "[The ex-convict] tied [Dodge] up because [Dodge] wasn't co-operating.")

   - 1 and 3 taken together alter whether or not Subject(S1)=Subject(SY).
   - 2 and 3 taken together alter whether or not Subject(SX)=Subject(SY).
   - Because the focusing algorithm prefers constancy in Subject Focus history, these alternations help us decide whether S1 or SX is more important in resolving pronouns in SY.

4. Whether Subject(S3) is the ex-convict or Dodge.[12] ("Then [the ex-convict] was arrested by the police" vs. "Then [Dodge] started screaming for help.")

   - This parameter (in conjunction with the others) determines whether Subject(S3)=Subject(SX) or Subject(S3)=Subject(SY) or neither (e.g., Subject(S3)=Direct-Object(SY)).[13]
   - Similar to 3 above, this parameter helps us determine how the NPs in S2 affect the resolution of the pronoun in S3.

5. Whether Subject(SX) was pronominalized.

6. Whether Direct-Object(SX) was pronominalized.

7. Whether Subject(SY) was pronominalized.

   - Parameters 5-7 help check how various patterns of pronominalization might affect the processing.

   By generating texts for all combinations of different values of these parameters, we are able to control for the influence of each focusing factor. Essentially, taken together, the texts capture all different grammatical roles, focus history patterns, and patterns of pronominalization.

---

12 Recall that when Subject(S3) is pronominalized, the referent of Subject(S3) is determined by the semantic slanting of the text.
13 There must be at least three entities in the discourse for Subject(S3) to cospecify neither Subject(SX) nor Subject(SY), otherwise Direct-Object(SY) is Subject(SX) or Subject(SY). Such a discourse does not appear in this paper.

The result of this procedure is a set of texts that can be presented to native speakers for judgments. The idea is that in texts that are judged to be acceptable, the reader sees no conflicts, so the focusing factors (grammatical role, pronoun history, and focus history) must agree with the semantic slanting. In texts that are judged unacceptable or ambiguous, there must be a conflict between these two sets of factors. Thus, gathering such judgments should identify how focusing factors and semantic factors influence how readers interpret pronouns, and on the basis of this information from the judgments, a possible extension of the focusing algorithm can be identified.

6.4 Finding an Extension Based on Judgments

To see how this methodology can be used to find an extension for a focusing framework, let us concentrate on determining how RAFT/RAPR should compute the focusing data structures for an "SX because SY" sentence in order to correctly resolve a subject pronoun in a simple sentence following an "SX because SY" sentence. In considering this question, recall that RAFT/RAPR prefers to resolve a subject pronoun in a simple sentence with the Subject Focus of the previous sentence. Thus, the specific question we need to address is how to compute the Subject Focus of an "SX because SY" sentence.[14]

Again, the factors that might determine how to compute the subject focus of a sentence are:

1. Syntactic form. For example, perhaps the Subject Focus should be computed as Subject(SX) (or, alternately, Subject(SY)), regardless of other factors. This would be appropriate if readers prefer to resolve the subject of the subsequent sentence so that it cospecifies Subject(SX) (or, alternately Subject(SY)), regardless of other focusing factors.

2. Pronominalization. For example, perhaps the Subject Focus should be computed as Subject(SX) or as Subject(SY) depending on which is pronominalized, regardless of other factors (unless both or neither are pronominalized). This would be appropriate if readers prefer to resolve the subject of the subsequent sentence on the basis of whether Subject(SX) or the Subject(SY) was pronominalized.

3. Focus history. For example, perhaps Subject Focus should be computed as Subject(SX) or Subject(SY) depending on which cospecified the Subject Focus of the sentence preceding the "SX because SY" sentence. This would be appropriate if readers prefer to resolve the subject of the sentence following the "SX because SY" sentence so it cospecifies the Subject Focus of the sentence preceding the "SX because SY" sentence.

4. Perhaps the interaction of two or more of the above factors (syntactic form, pronominalization, and focusing history) might influence how the Subject Focus should be computed.

6.4.1 Collecting SSD Judgments.

We contructed 32 discourses that manipulated the focusing factors as described above and pilot tested these on two of the authors and

---

14 For other focusing frameworks, we would be concerned with a different specific question. Note that the underlying question for each framework is one of how the data structures should be computed for an "SX because SY" sentence so that the referent of the subject in a (immediately) subsequent simple sentence can be computed in an efficient manner that is consistent with the processing of other simple sentences and other sentences that the framework can already handle.

185

two other people. This pilot group seemed to be primarily influenced by the syn-
tactic form of the sentences. In particular, their responses indicated a preference for
the subject focus of the "SX because SY" sentence to be the subject of SX. Because
asking subjects for judgments on such a large set of discourses was prohibitive, we
reduced the number of discourses to 13, including four that would critically examine
the hypothesis that the influence of syntactic form dominated the preferences. The
discussion below explains why the selected discourses test this hypothesis. The four
crucial discourses were distributed randomly throughout the test set.

Fifty-one participants were asked to judge the sentences as acceptable, awkward,
or ambiguous. (In some cases, a participant did not judge one of the discourses, so
there are instances with less than 51 judgments.) The participants selected were native
speakers of English working in linguistics or computational linguistics.

Suri (1993) performed a critical analysis of the experiment that explained why
other factors such as the infrequency of indefinite subjects in naturally occurring dis-
course, the use of passive or active voice, certain lexical choices, (potentially) stronger
reader identification with a victim/near-victim (or with a criminal), and order of text
presentation could not explain the distribution of judgments in our experiment. For
example, one may suspect that the use of the passive as in S1 of both Example 5 and
Example 6 may influence the judgments given. However, the experiment contained
pairs of examples where passive sentences were used for S1 and one was judged ac-
ceptable and the other unacceptable by the participants. Similar pairs of examples with
S1 in active voice were included with the same results. Thus it must be the case that
the judgments given were attributable to some factor other than active/passive voice.
Similar argumentation can be used to explain why the other factors did not influence
the judgments as well.

In the results reported below, judgments of "awkward" and "ambiguous" are
both considered to be "unacceptable." Since we want to derive an extension based
on what usages are acceptable, the distinction between awkward and ambiguous is
not relevant, since neither is acceptable. Table 1 contains the results of the four crucial
discourses contained in our experiment. In the table, the column labeled *Consensus*
indicates whether the majority of the participants judged the discourse acceptable or
unacceptable. The final column tells the statistical significance of this consensus, as
computed by a χ-square test. This test is computed separately for each discourse. In
order to compute χ-square, we must define the expected distribution of responses over
the categories. To get this, first, the overall distribution of the three categories can be
computed by summing the number of occurrences of each over the four discourses.
As can be seen from Table 1 (by summing each judgment column), roughly half the
judgments were "acceptable," a third were "awkward," and a sixth were "ambiguous."
These ratios can be used to define the expected value for each of the discourses. For
example, in Discourse 2, there were 50 responses. So, the expected number (according
to the overall distribution) of "acceptable" responses is about half the number of
responses, or 25. Using these expected frequencies, χ-square is then computed.[15]

For each discourse, a significant (p < .001) distribution of categories was demon-
strated. As is shown in Table 1, the group judged each discourse to be acceptable or
unacceptable with a high degree of significance. This indicates that the distribution of
judgments is very different from the random distribution that would result if partici-

---

15 An alternative method of computing expected frequencies is to make no suppositions regarding the
   three categories; that is, the expected frequency for each of the three categories is simply one third of
   the number of judgments for that discourse. With this method, the results are unchanged for three of
   the discourses, and Discourse 4 is significant at the .05 level.

Table 1
Acceptability judgments for the SSD's.

| | Number of Judgments | | | | |
| | | Unacceptable | | | |
Discourse	Acceptable	Awkward	Ambiguous	Consensus	Significance Level
1 (Ex. 5)	47	3	1	Acceptable	.001
2 (Ex. 6)	8	30	12	Unacceptable	.001
3	41	9	1	Acceptable	.001
4	9	25	16	Unacceptable	.001

pants had no real preferences. In intuitive terms, $p < .001$ for a discourse means that the majority category can indeed be called a consensus, since it is extremely unlikely that such a distribution would arise by chance.

6.4.2 Results of Collecting SSD Judgments. Our findings indicate that the syntactic form alone seems to most greatly influence what should be chosen as the Subject Focus of an "SX because SY" sentence (Suri 1993; Suri and McCoy 1993). In particular, we found that readers prefer to resolve the subject of a sentence following an "SX because SY" sentence so that the subject cospecifies Subject(SX), regardless of other focusing factors. We will refer to this finding as the **Prefer-SX Hypothesis**. This hypothesis indicates that RAFT/RAPR should be extended so that it computes the Subject Focus of an "SX because SY" sentence as Subject(SX). Consider how the judged discourses in Examples 5 and 6 support this conclusion.

Example 5
(S1)   Dodge was robbed by an ex-convict the other night.

(S2)   The ex-convict tied him up because he wasn't cooperating.

(S3)   Then he took all the money and ran.

Notice that in the discourse in Example 5 the semantic slanting should lead to the interpretation of S3 as "Then [the ex-convict] took all the money and ran." Thus, in this text, the semantic slanting favors resolving Subject(S3) as Subject(SX), and thus it would favor computing the Subject Focus of S2 to be Subject(SX). However, notice that the other focusing factors in the text all favor the Subject(SY) (i.e., Dodge) to be the subject focus of the complex clause. For instance, pronominalization favors computing the Subject Focus of S2 (which will be used to resolve Subject(S3)) to be Subject(SY) since Subject(SY) is pronominalized in S2, but Subject(SX) is not. The focus history would also favor the reading in which Subject(S3) is Dodge (i.e., corefers with the Subject(SY)) since Dodge was the Subject Focus of S1 and Dodge occurs in S2 as a subject.

Even though only the factor of syntactic form favors the interpretation that agrees with semantic slanting, of the 51 subjects in our experiment, 92% judged this discourse as acceptable. This supports the hypothesis that the syntax is the most important focusing factor and that it favors resolving a subject in a simple sentence following an "SX because SY" sentence so that the subject cospecifies Subject(SX).

The above hypothesis is further supported by the judgments given on a second discourse:

Example 6

(S1)   Dodge was robbed by an ex-convict the other night.

(S2)   The ex-convict tied him up because he wasn't cooperating.

(S3)   # Then he started screaming for help.[16]

In Example 6, one would expect subjects to judge the discourse to be "awkward" or "ambiguous" if the syntax factor overrides other focusing factors and causes Subject(SX) to be the preferred referent for the subject of the subsequent sentence. This is because the semantic slanting should lead to the interpretation of S3 as "Then [Dodge] started screaming for help," i.e., an interpretation for which Subject(S3)$\neq$Subject(SX). In fact, most subjects (82%), did judge the discourse as awkward or ambiguous. As a result, we have labeled (S3) with a "#", which traditionally denotes pragmatic ill-formedness.

To reiterate, in Example 6:

- On the basis of semantic slanting, Subject(S3)=Subject(SY) and Subject(S3)$\neq$Subject(SX).

- Subject Focus history would favor Subject(SY) for the Subject Focus of S2 since Subject(SY)=Subject(S1) (i.e., the Subject Focus of S1).

- Pronominalization would favor Subject(SY) for the Subject Focus of S2 since the referent of Subject(SY) is pronominalized in S2, but Subject(SX) is not.

- The discourse was judged "awkward" or "ambiguous."

Contrast this with Example 5 where the focus history and pronominalization favor computing the Subject Focus of S2 to be Subject(SY), the semantic slanting indicates Subject(S3) is Subject(SX), and the text was judged "acceptable."

Taken together, these judgments suggest that the reader prefers to resolve Subject(S3) with Subject(SX) regardless of the other focusing factors since this would explain the judgments as follows: In Example 6, the interpretation of Subject(S3) indicated by this focusing preference would be at odds with the interpretation forced by the semantic slanting and the text was judged awkward; in Example 5, the interpretation indicated by this focusing preference would agree with the interpretation forced by the semantic slanting and the text was judged acceptable.

The appropriateness of this conclusion should be verified through a corpus analysis. We discuss how such analyses should be performed in Section 6.5, and present preliminary results of such an analysis.

6.5 Confirming Preferences and Extensions through Corpus Analyses

Recall that the second part of our methodology involves using a corpus analysis to verify the findings of the SSD phase. In this section, we explain our methodology for verifying those findings by describing how we examined naturally occurring discourse

---

16 In all of the discourses, S3 began with *then*. This word may indeed be influencing the experimental results; perhaps the complex sentence type should be considered "SX because SY. Then...", and discourses without *then* need to be studied in a separate SSD experiment. A preliminary such study asking subjects for judgments on parallel texts with and without the *then* is included in Suri (1993). While the judgments of discourses without the *then* tended to agree with those with the *then*, the participants did not judge these two types of discourse identically. This suggests that the *then* is influencing the focusing preferences and further suggests that the *because* sentences with the *then* should indeed be considered as a different type from those without the *then*.

Table 2
Data for a subject in S_{n+1} cospecifying an SX subject (but not an SY subject).

An S_{n+1} subject pronoun cospecifies a subject in SX:	29
An S_{n+1} subject nonpronoun cospecifies a subject in SX:	5

Table 3
Data for a subject in S_{n+1} cospecifying an SY subject (but not an SX subject).

An S_{n+1} subject pronoun cospecifies a subject in SY:	13
An S_{n+1} subject nonpronoun cospecifies a subject in SY:	6

Table 4
Data for a subject in S_{n+1} cospecifying an SX and an SY subject.

An S_{n+1} subject pronoun cospecifies a subject in SX and a subject in SY:	22
An S_{n+1} subject nonpronoun cospecifies a subject in SX and a subject in SY:	5

to see how well the Prefer-SX hypothesis (proposed on the basis of the Semantically Slanted Discourses experiment described above) predicted the correct referents of pronouns in a sentence following a sentence of the form "SX because SY."

In the corpus analysis described below, we did not restrict our study of the corpus to cases where SX and SY were each a single clause, and thus for the discussion in this section, Section 6.5, SX does not necessarily denote a single clause, and likewise for SY. We chose to consider cases where SX or SY was not a single clause because the number of candidates of the appropriate form was quite small and we did not want to further limit the number of sentences that we could include in this corpus analysis. For each example in the corpus, one author coded the discourse, identifying the referents of the subject of SX, subject of SY, and subject of S3, and another author checked the coding to make sure it was reasonable.

We concentrated our analysis on studying the resolution of a subject in a sentence following a sentence of the form "SX because SY." We analyzed 81 text sequences where each text sequence contained an "SX because SY" sentence, S_n, followed by another sentence, S_{n+1}, which had one or more subjects that was/were coreferential with at least one subject in S_n. We were interested in which subject element(s) in the "SX because SY" sentence was/were being referred to by the subject(s) in the S_{n+1} sentence. From this analysis, we created three tables that captured the coreference relationships that we were interested in. We looked at how often a subject in S_{n+1} cospecified a subject of SX but not a subject of SY (and whether or not it occurred as a pronoun in S_{n+1}); this data is shown in Table 2. We also looked at how often a subject in S_{n+1} cospecified a subject of SY but not a subject of SX (and whether or not it occurred as a pronoun in S_{n+1}); this data is shown in Table 3. Finally, we looked at how often a subject in S_{n+1} cospecified a subject of SX *and* a subject of SY (and whether or not it occurred as a pronoun in S_{n+1}); this data is shown in Table 4.

We make the following comments:

- The corpora used for this study were the Brown corpus and several works of twentieth century literature that are in the public domain and are available on-line.[17]

---

17 At Project Gutenberg: http://www.gutenberg.org

- The information in Tables 2 to 4 is based on an analysis of 81 sentences (from the selected corpora) involving an "SX because SY" structure. We initially extracted more of such sentences from the corpora. We did not analyze some of these sentences because they involved biblical writing (and thus a very different style of writing), and we did not analyze others because they could not be classified in a straightforward fashion (because they involved multiple levels of complexity, or raised too many questions about how to code the sentence, etc.). Thus, this analysis is based on 81 sentences.

- In some cases an NP in S_{n+1} that cospecified an element of S_n also cospecified an S_{n+1} element that was prior to the NP in S_{n+1}. Thus, the interpretation of the NP might be based on the S_{n+1} element, rather than the S_n element. These cases were still counted like other cases where an S_{n+1} element cospecified an S_n element.

- Because S_{n+1} might be complex (in many instances it was), we write "An S_{n+1} subject pronoun/nonpronoun cospecifies ... " rather than "The S_{n+1} subject pronoun/nonpronoun cospecifies ...." Because S_n might be complex (in many instances it was), we write "... cospecifies *a* subject ..." rather than "... cospecifies *the* subject ...."

While we believe that the data analyzed is not conclusive, we note that it did not appear to contradict our Prefer-SX Hypothesis. We observe the following:

- A subject in S_{n+1} was more significantly more likely (34 times versus 19 times, $\chi^2 = 4.24$, $p < .05$) to cospecify only an SX subject than to cospecify only an SY subject. This suggests that a subject in SX (for an "SX because SY" sentence) is more likely to be a subject of a subsequent sentence than a subject in SY.

- A *pronominal* subject in S_{n+1} was significantly more likely (29 times versus 13 times, $\chi^2 = 6.10$, $p < .05$) to cospecify only an SX subject than to cospecify only an SY subject. This suggests that a pronominal subject in a sentence following an "SX because SY" sentence is more likely to cospecify the SX subject rather than the SY subject.

- An S_{n+1} subject cospecifying an SX subject was more likely to be pronominalized (85% pronominalization rate) than an S_{n+1} subject cospecifying an SY subject (68% pronominalization rate).

6.6 Implications for Proposed Extension

As noted above in Section 6.4, the Prefer-SX Hypothesis would indicate that RAFT/RAPR should be extended so that it computes the Subject Focus of an "SX because SY" sentence as Subject(SX). Since this corpus analysis did not appear to contradict the hypothesis, it also did not appear to contradict the appropriateness of the proposed extension. We point this out to clarify how a corpus analysis can be used to confirm proposed extensions, as well as confirming hypothesis about readers' preferences for pronoun interpretation.

Also, we stress that the SSD Methodology and the hypothesis generated by applications of the methodology have implications for extending other focusing frameworks, not just RAFT/RAPR. (See Suri [1993] or Suri and McCoy [1994] for a discussion of the implications of the Prefer-SX Hypothesis for a pronoun resolution algorithm

such as Brennan, Friedman, and Pollard [1987] based on centering.) This is important since it entails that applications of the SSD Methodology have general implications for focusing research, not just for extending the RAFT/RAPR framework.

7. Future Work

The SSD methodology must be employed in order to derive extensions to each framework for every kind of complex sentence. For example, sentences of the form "Because SX, SY" and "SX although SY" may have different extensions, and therefore will have to be examined separately under this methodology. However, we would like to be able to predict the extension suggested by the first (SSD construction) part of the methodology. One way to arrive at these predictions is to use a feature-based hierarchy of the connectives that appear in the complex sentences, as in Knott and Mellish [1996]. Then, if we have already found the proper extension for one complexity, say "SX because SY" sentences, we may predict that any connective that is synonymous (in terms of features) with *because* will have the same extension (provided the connective is used in a syntactically analagous sentence structure). Thus, we have bypassed the first part of the methodology. The second step of verifying the extension through a corpus analysis will still have to be performed. If this extension cannot be verified, then the first step will have to be done anyway.

Suri's (1993) critical analysis suggested the need to use the SSD Methodology to test the role of *then* and aspect in the interpretation of pronouns in the SSD experiment.[18] We stress that (because of the all the problems associated with corpus analyses that have been discussed in this paper) there is no clear way to test the influence of *then*, the duration of events, or aspect using only a corpus analysis.

As explained in Suri (1993), our claim that aspectual classification might play a role in the interpretation of pronouns in a subsequent sentence is distinct from the hypotheses of other researchers about the role of tense and aspect (e.g., Nakhimovsky 1988; Reichman 1978) in pronoun interpretation. In short, we believe Reichman (1978) only intended that tense and aspect could signal discourse segment boundaries, and thus indirectly influence pronoun interpretation, while Nakhimovsky (1988) claims that a change in time scale, aspectual class, or other temporal characteristics could signal a new discourse segment, and thus, indirectly influence pronoun interpretation. We propose that aspectual classification might affect pronoun interpretation *within* a discourse segment (Suri 1993).

Suri also reviewed literature on NP1-biased and NP2-biased verbs (see Caramazza et al. [1977] or Suri [1993]) and the implications of this work for future analyses.

8. Conclusions

The notion of local focusing and its influence on pronoun resolution has been found useful in many aspects of NLP. However, previous work on local focusing has ignored complex sentences even though they are prevalent in naturally occurring text. The problem that we faced was one of determining a reasonable way to extend a focusing algorithm to handle these sentences. Previous methodology (i.e., using semantically neutral text) was too simplistic and nearly impossible to utilize. A solely corpus-based analysis was impossible because of the variety of a priori decisions that needed to be

---

18 We thank Susan Brennan for first raising the question of how *then* affected the judgments in this experiment. We note that Walker (1993) explored the role of *now* in centering.

made and because of the complexity of interaction among factors in naturally occurring discourses. This work presents a methodology that calls for the systematic construction of texts. It relies on the potential tension of semantic factors with focusing factors to identify possible extensions of a focusing framework to account for a particular kind of complex sentence. The methodology has been used to extend a focusing framework (RAFT/RAPR) to handle one type of complex sentence ("SX because SY" sentences).

We illustrated the SSD Methodology by 1) explaining how the first part of the methodology led to the Prefer-SX hypothesis (a hypothesis about how readers prefer to resolve subjects in a sentence following an "SX because SY" sentence); and 2) discussing a (scaled-down) preliminary corpus analysis we performed to test the validity of the Prefer-SX hypothesis. We relaxed the methodology for the purposes of the preliminary corpus analysis because of the enormous amount of work required to perform the corpus analysis properly and because we felt that the corpus analysis would reveal more issues to be addressed prior to a proper corpus analysis. As reported in this paper, the corpus analysis did not refute our Prefer-SX hypothesis, and it did indeed reveal more issues that need to be addressed in this type, or any other type, of corpus analysis to examine focusing methodologies.

Note that one very important point raised by the preliminary corpus analysis is that the numbers derived during a corpus analysis are prone to represent not just the influence of the focusing framework, but the influence of world knowledge, semantics, and other pragmatic factors. Therefore, corpus analyses must be considered with this in mind. Because the problems that arose during our preliminary corpus analysis will arise during any kind of corpus analysis, we believe that our SSD Methodology is important for deciding how to extend a given focusing framework, and for comparing two focusing frameworks.

In fact, we feel that a corpus analysis to verify findings based on the SSD Methodology should be withheld until the processing of many kinds of complex sentences using the SSD Methodology has been analyzed. It would be better still to have an inference tool with which to reject referents based on semantics, world knowledge, pragmatics, etc. This would allow us to apply focusing algorithms to a corpus and automate the comparison of focusing algorithms, by adding functions to track and compute the frequency information (like the frequency information shown in Tables 2 to 4) that is needed for a corpus analysis.

The reader should note that the preferences for pronoun resolution that we identified (i.e., the Prefer-SX Hypothesis) refute an assumption sometimes made by researchers regarding complex sentences: that the clauses of complex sentences can be processed in a strictly linear order. Our findings indicate that the appropriate contents of *the focusing data structures* after processing S2 should be much more heavily influenced by the SX clause (and not by the SY clause, as the previous assumption would require). We stress that these findings are relevant for other focusing frameworks, not just RAFT/RAPR, and they indicate the importance of studying complex sentences.

Furthermore, as explained in detail in Suri (1993), the SSD Methodology can also be used to *compare* local focusing frameworks. Thus, this methodology allows the study of focusing phenomena and algorithms related to focusing phenomena.

References
Brennan, Susan E. 1998. Reference and local structure in interactive discourse. In Marilyn A. Walker, Aravind K. Joshi, and Ellen F. Prince, editors, *Centering Theory in Discourse*. Oxford University Press.

Brennan, Susan E., Marilyn W. Friedman, and Carl J. Pollard. 1987. A centering approach to pronouns. In *Proceedings of the 25th Annual Meeting*, pages 155–162. Association for Computational Linguistics.

Caramazza, Alfonso, Ellen Grober, Catherine Garvey, and Jack Yates. 1977. Comprehension of anaphoric pronouns. *Journal of Verbal Learning and Verbal Behavior*, 16:601–609.

Carter, David. 1987. *Interpreting Anaphors in Natural Language Texts*. John Wiley and Sons, New York.

Dahl, Deborah. 1986. Focusing and reference resolution in PUNDIT. In *Proceedings of the 1986 National Conference on Artificial Intelligence*, pages 1,083–1,088, Philadelphia, PA, August.

Dahl, Deborah and Catherine N. Ball. 1990. Reference resolution in PUNDIT. Technical Report CAIT-SLS-9004, UNISYS, Paoli, PA, March.

Gordon, Peter C., Barbara J. Grosz, and Laura A. Gilliom. 1993. Pronouns, names, and the centering of attention in discourse. *Cognitive Science*, 17(3):311–347.

Grosz, Barbara J., Aravind K. Joshi, and Scott Weinstein. 1983. Providing a unified account of definite noun phrases in discourse. In *Proceedings of the 21st Annual Meeting*, pages 44–50, Cambridge, MA, June. Association for Computational Linguistics.

Grosz, Barbara J., Aravind K. Joshi, and Scott Weinstein. 1995. Centering: A framework for modeling the local coherence of discourse. *Computational Linguistics*, 21(2):203–225.

Hobbs, Jerry. 1976. Pronoun resolution. Technical Report 76-1, Department of Computer Science, City College, City University of New York.

Hobbs, Jerry. 1978. Resolving pronoun references. *Lingua*, 44:311–338.

Hoffman, Beryl. 1998. Word order, information structure, and centering in turkish. In Marilyn A. Walker, Aravind K. Joshi, and Ellen F. Prince, editors, *Centering Theory in Discourse*. Oxford University Press.

Hudson-D'Zmura, Susan and Michael K. Tanenhaus. 1998. Assigning antecedents to ambiguous pronouns: The role of the center of attention as the default asignment. In Marilyn A. Walker, Aravind K. Joshi, and Ellen F. Prince, editors, *Centering Theory in Discourse*. Oxford University Press.

Kameyama, Megumi. 1986. A property-sharing constraint in centering.

In *Proceedings of the 24th Annual Meeting*, pages 200–206, Cambridge, MA. Association for Computational Linguistics.

Kameyama, Megumi. 1998. Intrasentential centering: A case study. In Marilyn A. Walker, Aravind K. Joshi, and Ellen F. Prince, editors, *Centering Theory in Discourse*. Oxford University Press.

Kameyama, Megumi, Rebecca Passonneau, and Massimo Poesio. 1993. Temporal centering. In *Proceedings of the 31st Annual Meeting*, pages 70–77. Association for Computational Linguistics.

Knott, Alistair and Chris Mellish. 1996. A feature-based account of the relations signalled by sentence and clause connectives. *Language and Speech*, 39(2-3):143–183.

Lappin, Shalom and Herbert J. Leass. 1994. An algorithm for pronominal anaphora resolution. *Computational Linguistics*, 20(4):535–561.

Linson, Brian. 1993. A distributional analysis of discourse transitions. In *Proceedings of the Workshop on Centering Theory in Naturally Occurring Discourse*, Institute for Research in Cognitive Science, University of Pennsylvania.

Nakhimovsky, Alexander. 1988. Aspect, aspectual class, and the temporal structure of narrative. *Computational Linguistics*, 14(2):29–43.

Palmer, Martha S., Deborah A. Dahl, Rebecca J. Schiffman, Lynette Hirschman, Marcia Linebarger, and John Dowding. 1986. Recovering implicit information. In *Proceedings of the 24th Annual Meeting*, pages 10–19, June. Association for Computational Linguistics.

Reichman, Rachel. 1978. Conversational coherency. *Cognitive Science*, 2:283–327.

Sidner, Candace L. 1979. *Towards a Computational Theory of Definite Anaphora Comprehension in English Discourse*. Ph.D. thesis, MIT, June.

Strube, Michael. 1996. Processing complex sentences in the centering framework. In *Proceedings of the 34th Annual Meeting*, pages 378–380. Association for Computational Linguistics.

Strube, Michael. 1998. Never look back: An alternative to centering. In *Proceedings of COLING-ACL '98: 36th Annual Meeting of the Association for Computational Linguistics and 17th International Conference on Computational Linguistics*, Montreal, Quebec, Canada, pages 1,251–1,257.

Suri, Linda Z. 1993. *Extending Focusing Frameworks to Process Complex Sentences and to Correct the Written English of Proficient*

Signers of American Sign Language. Ph.D. thesis, University of Delaware. Available as Dept. of CIS Technical Report TR-94-21.

Suri, Linda Z. and Kathleen F. McCoy. 1993. *Focusing and Pronoun Resolution in Particular Kinds of Complex Sentences*. In *Proceedings of the Workshop on Centering Theory in Naturally Occurring Discourse*, Institute for Research in Cognitive Science, University of Pennsylvania.

Suri, Linda Z. and Kathleen F. McCoy. 1994. RAFT/RAPR and centering: A comparison and discussion of problems related to processing complex sentences. *Computational Linguistics*, 20(2).

Walker, Marilyn. 1989. Evaluating discourse processing algorithms. In *Proceedings of the 27th Annual Meeting*, pages 251–261. Association for Computational Linguistics.

Walker, Marilyn. 1993. *Initial Contexts and Shifting Centers*. In *Proceedings of the Workshop on Centering Theory in Naturally Occurring Discourse*, Institute for Research in Cognitive Science, University of Pennsylvania.

Walker, Marilyn, Masayo Iida, and Sharon Cote. 1994. Japanese discourse and the process of centering. *Computational Linguistics*, 20(2).