# Focusing the Comprehension of Definite
Anaphora -- Sidner

Anaphora -- pointing back, a discourse element which relies on a previous discourse element for its interpretation.

Problem -- determining the antecedent (the previous discourse element which is used in the interpretation of the anaphor)

The Real World

Hearer's Mind

Discourse Model

phrase67 (r1)

Anaphoric-NP2 (r1)
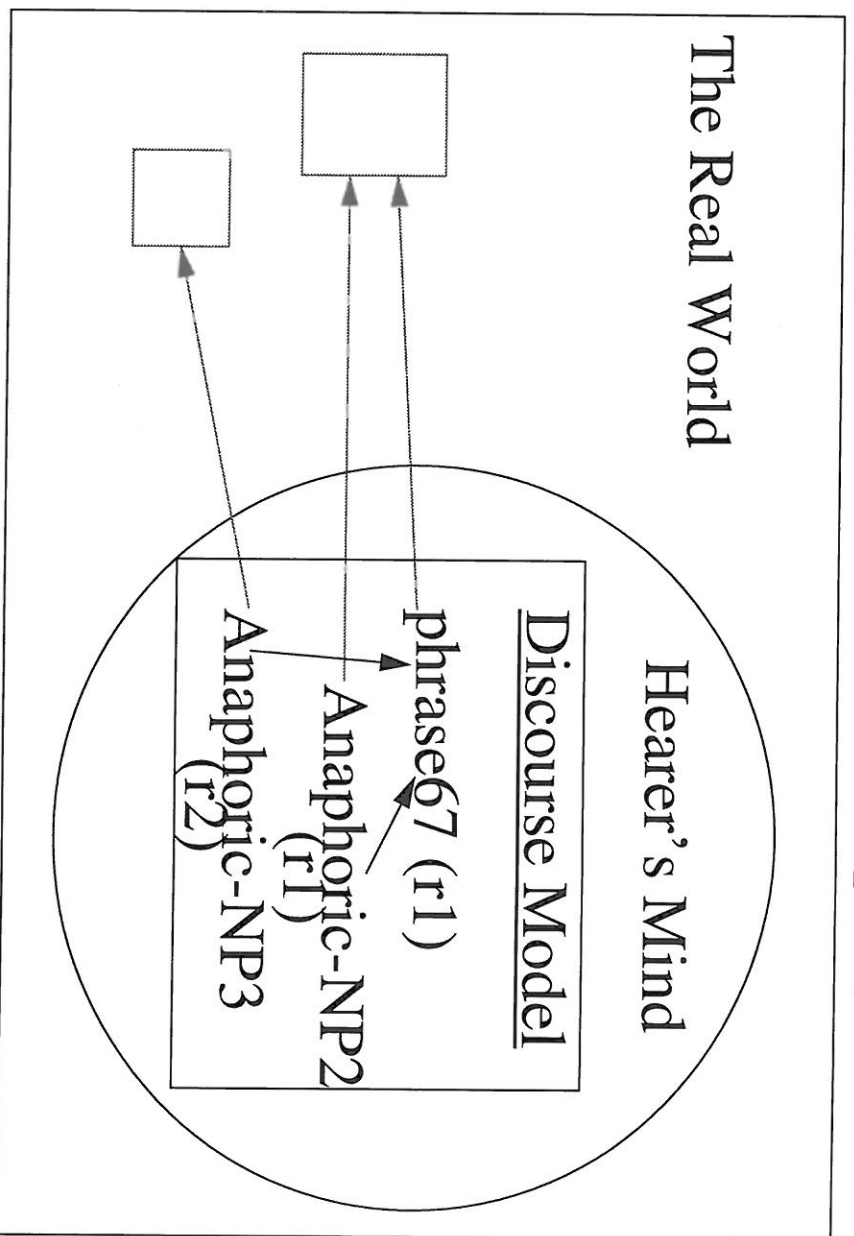
Anaphoric-NP3 (r2)

Anaphoric-NP2 and 3 co-specify phrase67.

Focusing

discourse -- a connected piece of text or spoken language of more than one sentence

focus -- the thing that is being currently centered on in a discourse

How do we know what is focused on?

1. special words indicate "sameness" of focus (in particular, anaphoric expressions)

2. assumed shared knowledge helps establish connections

Focusing information and Shared Knowledge can be used in determining co-specification relationships for anaphoric expressions

1. Focusing offers possible co-specificiers

2. Shared Knowledge is used to confirm or deny

Assumptions about nature of communication

1. Speaker is assumed to be communicating about something (that something is called the focus)

2. Speaker assumed that the hearer can identify the focus of the discourse

3. Speaker is not trying to confuse or deceive the hearer

4. Speaker assumes the hearer has certain knowledge about the real world which can be used to reason about referring expressions during the communication process

Model of Focusing

1. Focusing data structures are established for the initial (previous) sentence

2. A set of interpretation rules use the focusing data structures (with confirmation from world and shared knowledge) to interpret the anaphoric expressions in the current sentence.

3. The focusing data structures are updated based (in part) on the co-specification relationships of anaphors in the current sentence.

Data Structures Used in Tracking Focus

AF - Actor Focus - thematic agent of the sentence

PAFL - Potential Actor Focus List - contains all NP's that specify an animate element of the database and do not occur in agent position

CF - Current (discourse) focus - the thing sentence is most centrally about - generally the theme of the sentence

PFL - Potential Focus List - all other NP's (non CF or AF) in the sentence ordered by surface order, followed by the AF, followed by the VP

Focus Stack - when ever focus is changed, keep old foci on a stack

Options for (CF) focus moves:

- Continue talking about the same things
  $CF_i = CF_{i-1}$

- Talk about something just introduced
  $CF_i = $ Member of PFL_{i-1}

- Return to a topic of previous discussion
  $CF_i = $ Member of Focus Stack$_{i-1}$
  this options causes a popping of the
  elements on the stack between the top of
  stack and the element chosen as CF

Go through these in order, preferring to make CF
the first of these elements which were co-
specified (with a pronoun) in the current sentence

Anaphors are prioritized - pronouns better CF
indicators than are definite NPs

AF and CF should be different (if possible), so try
to avoid making $CF_i = AF_i$

Pronoun Resolution Rules -

- Pronoun in agent position

  1. the previous AF is chosen for its co-specification

  2. choose members of the PAFL in order

  3. choose CF (of previous sentence)

- Pronoun in non-agent position

  1. the previous CF

  2. choose members of the previous PFL in order

  3. choose top-most element from stack (and continue through stack in order)

Some Sample Discourses:

1. I want to schedule a meeting with George, Jim, Steve and Mike.

2. We can meet in my office.

3. It's kind of small,

4. but we'll only need it for about an hour.

---

1. Last week there were some nice strawberries in the refrigerator.

2. They came from our food co-op and were unusually fresh.

3. Later I discovered it was Mark who had eaten them.

4. Mark has a hollow leg, and it's impossible to keep food around when his stomach needs filling.

1. Jerome took his pigeon out on a leash.

2. Since he was trying to train it,

3. he hollered "heel" and "run" at it,

4. as they sauntered along.

1. I wanted to go to the movies on Saturday.

2. John said he'd come too, but Bill decided to stay home and study.

3. So we went and afterwards had a beer.

1. Mary took a nickel from her toy bank yesterday.

2. She put it on the table near Bob.

1. Sandy walked her dog near a bull one day.

2. He walked quietly along.

1. Sandy walked her dog near a bull.

2. She saw how he threw back his great menacing horns.

1. Cathy wants to have a big graduation party at her house.

2. She cleaned it up

3. so that there would be room for everyone.

Example:

"(S1) First, in summer I live at home with my parents.
(S2) I can budget money easily.
(S3) I did not spend lot of money at home because at home we have lot of good foods, I ate lot of foods.
(S4) While living at college I spend lot of money because — go out to eat almost everyday.
(S5) At home, sometimes my parents gave me some money right away when I need —.
(S6) While in college, I could not ask my parents for money right away because I live in Washington DC and my parents live in Illinois."

Focus Data Structures after S1:

CF	HOME
PFL	SUMMER, and the LIVE VP
AF	I
CF stack	empty
PFL stack	empty

Focus Data Structures after S2:

CF	I
PFL	MONEY, EASILY, and the BUDGET VP
AF	I
CF stack	HOME
PFL stack	SUMMER, and the LIVE VP

Focus Data Structures after S3:

CF	MONEY
PFL	HOME, NOT SPEND VP, and GOOD FOOD, HOME, and the HAVE VP.
AF	I
CF stack	I, HOME
PFL stack	PFL of S2, followed by the PFL stack of S2

Example:

"(S1) *First, in summer I live at home with my parents.*
(S2) *I can budget money easily.*
(S3) *I did not spend lot of money at home because at home we have lot of good foods, I ate lot of foods.*
(S4) *While living at college I spend lot of money because __ go out to eat almost everyday.*
(S5) *At home, sometimes my parents gave me some money right away when I need __.*
(S6) *While in college, I could not ask my parents for money right away because I live in Washington DC and my parents live in Illinois.*"

Focus Data Structures after S4:

CF	MONEY
PFL	COLLEGE, the SPEND VP, ALMOST EVERY DAY, the TO EAT VP, and the GO OUT TO EAT VP.
AF	I
CF stack	MONEY, I, HOME
PFL stack	PFL of S3, followed by the PFL stack of S3

Focus Data Structures after S5:

CF	MONEY
PFL	NEED VP, MY PARENTS, HOME, GIVE VP
AF	I
CF stack	MONEY, MONEY, I, HOME
PFL stack	PFL of S4, followed by the PFL stack of S4

Centering (and Pronoun Resolution) - Grosz, Joshi, & Weinstein 83 and Brennan, Friedman, and Pollard 87

The centers of a sentence serve to integrate the sentence into the discourse

The Data Structures:

backward-looking center or the *(Cb)*: this is intended to capture that item which ties the current utterance with the previous utterance in the discourse. (This is often referred to as *the center*.)

list of *forward-looking centers* *(Cf)*: the list of all NPs in the utterance. These elements can potentially be the Cb of the next utterance.

preferred center or *(Cp)*: this is simply the first element of the Cf (the subject of the utterance for English). This is the element that is considered most likely to become the Cb of the next utterance.

Compute the Cb using the Cb constraint:

Cb constraint: the Cb of utterance U_n, $Cb(U_n)$, is the highest-ranked element of the Cf of the previous utterance, $Cf(U_{n-1})$, that is realized in the current utterance.

Cf is ordered list of all NP's in sentence:

For English, Cf ordered: subject, object, object2, other subcategorized functions, and adjuncts.

1. Who did Max see yesterday? ---------------- cf = (Max yesterday)

2. Max saw Rosa. ------------------- cb = Max; cf = (Max Rosa)

1. Did anyone see Rosa yesterday? -- cf = (anyone Rosa yesterday)

2. Max saw Rosa. -------------- cb = Rosa; cf = (Max Rosa)

1. How is Rosa? ------------------- cf = (Rosa)

2. Did anyone see her yesterday? -- cb = Rosa; cf = (anyone Rosa yesterday)

3. Max saw her. --------------- cb = Rosa; cf = (Max Rosa)

Constraint on the Speakers:

If the Cb of the current utterance is the same as the Cb of the previous utterance, a pronoun should be used.

note: pronouns may be used for other entities, as long as one is used for the center. This is not a hard rule, it may be broken but then has some effect

1. John hasn't had much homework lately. -- cf = (John homework)

2. He has trouble leaving people alone when he isn't busy.
   cb = John, cf = (John people)

3. He called up Mike yesterday. -- cb = John, cf = (John Mike yesterday)

4. *He was annoyed by John's call. (he = Mike)

1. John hasn't had much homework lately.

2. He has trouble leaving people alone when he isn't busy.

3. He called up Mike yesterday.

4. ?He was studying for his driver's test.

5. ?He was annoyed by John's call.

A Problematic Example for Sidner -- oops -- note I gave wrong pronoun resolution rule for Sidner during last class!

1. I haven't seen Jeff for several days. --------------AF = I, CF = Jeff

2. Carl thinks he's studying for his exams.
   he = Jeff (because AF doesn't work, CF does) --------AF = Carl, CF = Jeff

3. But I think he went to the Cape with Linda.
   he = Carl (because prefer AF for pronoun)

Note correct interpretation would be given for:
He thinks he studies too much.

But most agree he = Jeff! Consider centering's treatment:

1. I haven't seen Jeff for several days. ------ Cf = (I Jeff)

2. Carl thinks he's studying for his exams.
   he = Jeff (because ?)

3. But I think he went to the Cape with Linda.
   Cb = Jeff; Cf = (I Jeff...)
   Because only 1 pro = Cb previous sentence

Note: they say they can handle 2 pro case, but they don't say how they resolve pronouns!

1. Susan drives a Ferrari. ------ Cf=(Susan, Ferrari)
2. She drives too fast. ------ Cb=Susan, Cf=(Susan)
3. Lyn races her on weekends. --- Cb=Susan, Cf=(Lyn, Susan)
4. She often beats her.

3. There are two possible interpretations:

Case 1: For the interpretation of U3 as ``Lyn races [Susan] on weekends," we would compute the Cf(U3) to be ([Lyn], [Susan]), and thus Cp(U3)=[Lyn]. The Cb(U3) would be [Susan] since [Susan] would be the highest-ranked element (in fact, the only element) of the previous forward-looking centers list that is realized in the current utterance. Thus, under this interpretation, the utterance would result in a *retain* transition.

Case 2: For the interpretation of U3 as ``Lyn races [Lyn] on weekends," there would be no Cb since nothing in U3 would realize anything in Cf(U2). The Cp would be [Lyn] since the referent of the subject of U3 would be [Lyn]. However, this interpretation is rejected by filtering of contraindices.

---

4. There are two possibilities (once contraindexing is considered):

Case 1: For the interpretation, ``[Lyn] often beats [Susan]," we would compute the Cf(U4) to be ([Lyn], [Susan], [Lyn]), and thus Cp(U4)=[Lyn], and the Cb(U4) would be [Lyn] since that would be the highest-ranked element of Cf(U3) realized in U4. Therefore, Cb(U4) neq Cb(U3), and Cb(U4)=Cp(U4), making this interpretation result in a *smooth-shift* transition.

Case 2: For the interpretation, ``[Susan] often beats [Lyn]," we would compute the Cf(U4) to be ([Susan], [Lyn]), and Cp(U4)=[Susan], and the Cb(U4) would be [Lyn] since that would be the highest-ranked element of Cf(U3) realized in U4. Therefore, Cb(U4) neq Cb(U3), and Cb(U4) neq Cp(U4), making this interpretation result in a *rough-shift* transition.

1. Brennan drives an Alfa Romeo.
2. She drives too fast.
3. Friedman races her on weekends.
4. She often beats her.

Some other examples:

1. Carl works at HP on the Natural Language Project.
   Cb = Carl; Cf: (Carl, HP, NatLang); Continuing
2. He manages Lyn.
   Cb = Carl; Cf: (Carl, Lyn); he=Carl; Continuing
3. He promised to get her a raise.
   Cb = Carl; Cf: (Carl, Lyn, raise); he=Carl; Continuing
4. She doesn't believe him.
   Cb=Carl; Cf: (Lyn, Carl); she=Lyn; him=Carl; Retaining

1. Who is Max waiting for?
2. He is waiting for Fred.
3. He invited him to dinner.

---

1. Brennan drives an Alfa Romeo.
2. She drives too fast.
3. Friedman races her on weekends.
4. She goes to Laguna Seca.
   She =?

RAFT/RAPR and Extending Focusing Frameworks -- Suri 93

Data Structures

Current Focus (CF): the item computed to be the local focus of the sentence.

Potential Focus List (PFL): all NPs other than the CF and SF, ordered according to the following: direct object, indirect object, all other NPs in surface order within the clause.

Subject Focus (SF): basically, the surface subject of the clause.

Potential Subject Focus List (PSFL): all NPs other than the SF and CF, ordered as follows: direct object, indirect object, all other NPs in surface order within the clause.

CF stack, SF stack, PFL stack, PSFL stack. We stack the foci and foci lists after each sentence.

SF and CF Computation

- SF is surface subject of sentence.
- Initial CF computed on basis of Syntax.
- After the first sentence, writer's options:

1. Continue talking about the same thing. In this case, the CF does not change.

2. Talk about something just introduced. In this case, the CF becomes a member of the previous sentence's PFL or the SF.

3. Return to a local focus of previous discussion. In this case, that item must have been the CF of a previous sentence.

4. Discuss something mentioned earlier, that was not the local focus. In this case, the item will be on the PFL stack or the SF stack.

5. Discuss something implicitly related to the last CF, an element just mentioned in the previous sentence, an old topic, an old SF, or something just mentioned earlier.

Decision as to which move was taken based on:

- the syntactic form of the current sentence
- the grammatical roles of the NPs of the current sentence
- whether the NPs co-specify elements previously mentioned in the discourse

and, if so,

- whether (the referent of) each NP co-specifies the CF of the previous sentence
- the SF of a previous sentence
- a member of the PFL of previous sentence
- an element of a stacked focus data structure.

Pronoun Resolution Rules

non-subject third person singular pronoun

1. the CF (of the last sentence)
2. the SF,
3. members of the PFL
4. the members of the CF stack, SF stack, PFL stack and PSFL stack.

subject third person singular pronoun

1. the SF
2. the CF,
3. members of the PSFL
4. the stacked elements.

CF Computation (in Simple Sentences) is based on the following interacting criteria:

Co-specification: prefer elements that co-specify an element in a focusing data structure over elements just introduced.

The type of realization of each element: prefer NPs realized as pronouns over those realized with full NPs.

Anaphoric status: prefer anaphors to non-anaphors.

Which focusing data structure is co-specified by each NP (that co-specifies something in previous text).
CF > last PFL or last SF > Stacks

Syntax: we prefer for the CF to be a non-subject rather than a subject, although the CF can be the subject.

Syntactic forms and clue words: e.g., there-insertion, ``but'' influence CF choice

---

Example of Pronoun Resolution with RAFT/RAPR

(S1) Susan drives a Ferrari
SF= [Susan],
CF=[Ferrari] (based on syntax)

(S2) She drives too fast.
she=Susan,
SF= [Susan],
CF=[Susan] %\\

(S3) Lyn races her on weekends.
her=Susan
SF= [Lyn],
CF=[Susan],
PFL=[weekends]

(S4) She often beats her.

In (S4), we first try to resolve ``She'' using the SF and ``her'' using the CF. This interpretation is not rejected on the basis of inferencing with semantic factors or syntactic information.

Why Stack? (Centering doesn't):

She left and *he* sat behind the old desk, one that looked fine and would have suited *him* except that it was considered used and therefore not good enough for a new lawyer at Bendini, Lambert & Locke.

The office was fifteen by fifteen, with two six-foot windows facing north and staring directly into the second floor of the old building next door.

Not much of a view.

With a strain, *he* could see a glimpse of the river to the northwest.

The Firm, p. 65

Note: pronoun referent to an element from two sentences back (and no indication of a discourse segment pop).

Another Stacking Example (from close to home!)

On November 20 John Hopkins will be hosting the 7th Maryland ``Theory Day'', which several of us will be attending.

There will be no meeting of SIGALG on that day.

Particularly students interested in algorithms are encouraged to attend.

Errol Lloyd, 11/92, electronic newsgroup posting

Notice that the students are encouraged to attend the meeting mentioned in the first sentence, not the meeting in the second sentence.

Comparing Centering and RAFT/RAPR

Stacking Focus Information

A major difference between the two frameworks is that by maintaining a stack of CFs, RAFT/RAPR allows a writer to return with a pronoun to a topic of a previous sentence. Centering, which has no counterpart to our stacks, does not.

Consider:

With a strain, *he* could see a glimpse of the river to the northwest.

The walls were Sheetrock and bare.

She *had* picked out some artwork.

He determined that the Ego Wall would face the desk, behind the wing chairs.

(*The Firm*, p. 139)

Discourse-Initial Sentences

Centering can make no predictions about pronouns until the third sentence in a discourse (segment). Consider the following:

1. Lyn races Susan on weekends.
2. She races Jack during the week.

One way around this problem is for a likely Cb (e.g., the subject?) for the first sentence to be calculated and used in pronoun resolution.

But:

1. Lyn races Susan on weekends.
2. Jack races her during the week.

Problem: Centering does not take grammatical role into account -- 1 pronoun MUST be Cb of previous sentence.

Pronominalization and Grammatical Roles

1. Susan drives a Ferrari.
2. She drives too fast.
3. Lyn races her on weekends.
4. She wins a lot of trophies.
(RAFT/RAPR: She=[Lyn], centering: She=[Susan])

Consider what happens when the ``her'' in 3 is replaced by a full NP:

1. Susan drives a Ferrari.
SF=[Susan], CF=[Ferrari];
Cf=([Susan], [Ferrari]
2. She drives too fast.
SF=[Susan], CF=[Susan];
Cb=[Susan], Cf=([Susan])
3. Lyn races Susan on weekends.
SF=[Susan], CF=[Susan];
Cb=[Susan], Cf=([Susan])
4. She wins a lot of trophies.
SF=[Lyn], CF=[Susan];
Cb=[Susan], Cf=([Lyn], [Susan])
(RAFT/RAPR: ``She''=[Lyn];
centering: ``She''=[Susan])

Perhaps 3 is the start of a new discourse segment?

• then centering has to tell us how to find discourse segments

• centering has to handle discourse segment initial sentences

Consider the same initial discourse, with a different 4th sentence:

1. Susan drives a Ferrari.
SF=[Susan], CF=[Ferrari];
Cf=([Susan], [Ferrari]
2. She drives too fast.
SF=[Susan], CF=[Susan];
Cb=[Susan], Cf=([Susan])
3. Lyn races Susan on weekends.
SF=[Susan], CF=[Susan];
Cb=[Susan], Cf=([Susan])
4. Jack races her during the week.
SF=[Lyn], CF=[Susan];
Cb=[Susan], Cf=([Lyn], [Susan])
centering: ``her''=[Susan];
(RAFT/RAPR: ``her''=[Susan])

Comparing the Frameworks (Simple Sentences)

1. RAFT/RAPR maintains two foci: a subject focus and a current focus. Centering maintains one focus, the Cb.

2. RAFT/RAPR resolves non-subject pronouns in a different manner than subject pronouns, while centering does not.

3. RAFT/RAPR resolves pronouns by searching data structures in an order based on several factors, including the importance of grammatical roles, and preferences for focusing movement. Centering resolves pronouns by generating all possible co-specifications and then filtering and ranking them based on a number of constraints and rules.

---

4. By using Rule 1 to eliminate possible co-specifiers, centering mixes the process of pronoun resolution with focus computation. The RAFT/RAPR approach resolves pronouns and then updates the focus.

5. Because the *abstract* preferences *underlying* the frameworks share much in common, the approaches very often make the same predictions.

6. RAFT/RAPR presents possible referents for pronouns one possibility at a time, and if pragmatic, semantic and general knowledge inferencing rejects a referent, RAFT/RAPR proposes an alternative. Centering, on the other hand, (in addition to requiring this same kind of inferencing) sometimes suggests multiple possibilities for the co-specifications of pronouns in a sentence.

Import of Above Analysis

Two observations:

1. it is very difficult to identify the focus of a discourse-initial sentence and neither framework is likely to process all discourse-initial text correctly

   This problem is compounded for centering since it works on discourse *segments*.

2. All we have done is shown a couple of problematic examples -- clearly this is insufficient!

   The question that must be addressed is which framework would correctly process the most *naturally occurring* discourse-initial texts.

Need to do a Corpus Analysis

Two things we want to do:

1. Compare two different focusing algorithms

2. Extend our focusing algorithm to handle complex sentences

A Corpus analysis seems most reasonable! But....

- No one has specified how to handle complex sentences in their algorithms
  - Complex sentences are very prevalent
  - The handling of such sentences are crucial for determining how the algorithm should be applied
    - Sidner (for example) broke up complex sentences differently in different places
  - Some arguments (e.g., used by Sidner and Centering) rely on complex sentences, but they are not explained. Consider: "Think" examples.

Basic Issues for Processing Complex Sentences

Central Question: Should the clauses of a complex sentence be processed in linear order, resolving pronouns and updating the focusing data structures as if one were processing a sequence of isolated simple sentences? Or is some other method more appropriate?

The answer to this question might differ for different types of complex sentences.

- "SX because SY"
- "SX but SY"
- "SX when SY"
- "NP₁ thinks SX"
- "NP₁ promised NP₂ VP"
- "NP₁ persuaded NP₂ VP"

Complex Sentences: Central Question

- What information should be used to propose referents for the pronouns *within* a complex sentence?

- What should the focusing data structures be *after* the complex sentence is processed?

Why can't we just analyze Constructed Semantically-Neutral Text?

Construct texts which exhibit precisely the questioned aspects, and gather judgements.

- For determining how to handle complex sentences, it became very difficult to construct texts that exhibited what we needed AND sounded natural.

- Difficult to isolate all factors that might be at work.

- How often (if ever) does a discourse (like the constructed one) actually occur?

- What other influences might not one account for?

Challenges with a Corpus Analysis

What to do?

- Take a possible extension (or one focusing framework)

- Count

  - number of correct pronoun resolutions?

  - number of referents considered before correct pronoun is happened upon?

  - (consider being WRONG vs taking time to get there)

  - number of times a pronoun is used when framework suggests it should be used?

  - number of ambiguous pronouns noted as ambiguous in framework?

Challenges of Corpus Analysis (Cont.)

- what do you count as a pronoun?

- How do you account for the influence of semantics?

- If trying to extend a framework, which frameworks do you test?

  - this method requires identifying the framework in advance

  - no framework can emerge as a result of the corpus analysis

Doing a corpus analysis requires already knowing how to handle each kind of complex sentence!

Semantically Slanted Discourse Methodology

- Construct discourses that are intentionally loaded or slanted for pronoun interpretation based on world knowledge, other pragmatic factors, and semantics.

- We contend that in a semantically-slanted discourse, if the text seems awkward, or one needs to re-interpret a pronoun, then the focusing and syntactic preferences for pronoun resolution are at odds with the preferences based on semantics, other pragmatic factors, or world knowledge.

- On the other hand, if the text seems acceptable/natural, then we contend the preferences for pronoun resolution based on focusing agree with preferences based on semantic-slanting.

- Thus, gathering acceptability judgments about semantically-slanted discourses should help us identify what the focusing preferences are.

Influences on Pronoun Resolution:

Focusing Factors (Part of the Framework)

- whether an NP co-specifies an element in one of the focusing data structures, or introduces a new element

- the type of each co-specifier, i.e., a full NP vs. pronoun

- which focusing data structure is co-specified by each co-specifier

- grammatical roles of elements in the sentence

- syntactic form (e.g., the use of a \em there]-insertion structure), and clue words (e.g., ``but'')

- how complex sentence is segmented into utterances.

Non-Focusing Factors: semantics, world knowledge, pragmatics

Semantically Slanted Discourse Methodology (Background)

1. Use a particular kind of constructed discourse to allow an extension of a focusing framework to emerge.

2. Test that extension using a corpus analysis.

Constructing Discourses:

We construct discourses of the following form, for which the interpretation of NPs in S3 is fully determined by the semantics of the text and world knowledge:

1. Simple-sentence

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

3. Simple-sentence

This we can examine influence of complexity on:

- resolving pronouns in S2 (in particular, in the non-initial clause)

- updating the focusing data structures after S2 so that the pronouns of S3 can be correctly resolved

```
\input chapter4/sxbcsy
\subsubsection{Using the SSD Methodology}\label{how-to-use-it}
```

Looking at "SX because SY" Sentences - - Subject Pronouns

We examined discourses that are "variations" of:

1. Dodge was nearly robbed by an ex-convict the other night.
2. [Dodge] captured [the ex-con] because [the ex-con] was so stupid and clumsy.
3. Then [Dodge] called the police.

We needed to construct variations of this text in order to tease out how the various focusing factors interact.

Some questions requiring variations:

1. How should subject(SY) be resolved?
   - should the algorithm prefer that it co-refer with subject(S1) or subject(SX)?
2. How should subject(S3) be resolved?
   - Preferring subject(SX) always?
   - Preferring subject(SY) always?
   - Preferring subject(SX or SY) depending on which is pronominalized?
   - Preferring subject(SX or SY) depending on which is co-referential with subject(S1)?

In order to answer these questions, we must make up a number of texts which vary these factors.

Text variations parameters:

1. Whether Subject(S1) is the ex-convict or Dodge.
2. Whether Subject(SX) of S2 is the ex-convict or Dodge.
3. Whether Subject(SY) of S2 is the ex-convict or Dodge.
4. Whether Subject(S3) is the ex-con or Dodge.
5. Whether Subject(SX) was pronominalized.
6. Whether DirectObject(SX) was pronominalized.
7. Whether Subject(SY) was pronominalized.

These combinations controlled:

- Parameter 1 controls for the influence of SF history on computing SF(S2) (and thus indirectly on resolving the subject pronoun in S3).
- Parameter 2 controls for the influence of the Subject(SX) on computing SF(S2) (and resolving pronouns in S3).
- Parameter 3 controls for the influence of the Subject(SY) on computing SF(S2).
- In combination, parameters 1--3 control for whether readers prefer to resolve the Subject(SY) with Subject(SX) or SF(S1).
- In combination, parameters 2--4 determine whether Subject(S3) co-specifies Subject(SX) and/or Subject(SY). Thus, the text variations based on these parameters help determine what the Subject Focus of S2 should be.

- We use parameter 4 to control for empathy or for a reader's discourse-independent preferences for a pronoun to refer to the ex-convict or Dodge. \Item In combination, parameters 5 and 6 control for the influence of pronominalization in S2 on resolving pronouns in SY and in S3.

- In combination, parameters 5-7 control for the influence of pronominalization in S2 on resolving pronouns in S3.

- Parameters 2 and 3 control whether Subject(SX) refers to the same element as Subject(SY). This control is important for testing the role of the Subject(SX) and the role of the Subject(SY) in pronoun resolution of Subject(S3).

- Parameters 2 and 3 also control for the influence of the syntactic roles (Subject versus Direct Object) of elements of S2 in resolving pronouns in S3. Recall that DirectObject(SX) neq Subject(SX).

Prefer SX Hypothesis:

- Readers prefer to resolve Subject(SX) with Subject(S1).

- Readers prefer to resolve Subject(SY) with Subject(SX) (rather than Subject(S1)).

- Readers prefer to resolve Subject(S3) with Subject(SX).

Test this with two kinds of discourses:

1. Discourses where the semantic-slanting determined that Subject(S3) co-specified Subject(SX), BUT focusing factors favored Subject(SY)
Here: acceptable would support

2. Discourses where semantic-slanting AND focusing factors determined Subject(S3) co-specified Subject(SY)
Here: awkward or ambiguous would support

SHOW SAMPLES AND RESULTS