

Resolving Pronominal Reference to Abstract Entities

Donna K. Byron

Department of Computer Science

P.O. Box 270226

Rochester, NY 14627

dbyron@cs.rochester.edu

Abstract

This paper describes PHORA, a technique for resolving pronominal reference to either individual or abstract entities. It defines processes for evoking abstract referents from discourse and for resolving both demonstrative and personal pronouns. It successfully interprets 72% of test pronouns, compared to 37% for a leading technique without these features.

1 Introduction

Automated techniques for pronoun interpretation have generally tried to resolve only those pronouns that corefer with another noun phrase (henceforth **NPC pronouns**). Choosing an antecedent based on the attentional state of the discourse works well for NPC pronouns; so for quite some time, pronoun resolution work has been dominated by salience-based techniques.

While a fruitful line of research for NPC pronouns, salience-based techniques do not work for pronouns whose antecedent is not a noun phrase. Such pronouns are common in spoken discourse. Eckert and Strube (2000) note that 22% of pronouns in a set of Switchboard dialogs had other-than-noun-phrase antecedents, and 33% had no antecedent. Byron and Allen (1998) report that only 50% of pronouns in the TRAINS93 corpus are NPC, and Botley (1996) found 20% of pronouns in his corpus to have no noun phrase antecedent. This data suggests that spoken dialog understanding systems must be able to interpret both NPC and non-NPC pronouns.

The PHORA pronoun resolution algorithm, implemented within the TRIPS spoken dialog

system at the University of Rochester, does just that. PHORA applies semantic filtering as a complement to salience calculations, which enables it to resolve pronouns referring to less salient *abstract* entities such as actions, propositions, and kinds. As a result, PHORA works equally as well on personal (e.g. it, him, them) as on demonstrative (e.g. this, those) pronouns. It is based on the observation that a pronoun that does not refer to the most salient item is typically constrained by its context to be incompatible with items more salient than the intended referent. Sidner (1983) describes this tradeoff between semantics and focus, but it has not been incorporated into pronoun resolution software until now. The details of PHORA will be described in Section 3; but first, Section 2 reviews relevant prior work that provided a basis for PHORA.

2 Background and Related Work

This section describes relevant prior work on anaphoric pronoun resolution. Anaphora resolution has been defined as a two-phase process: 1) identify what a text makes available for anaphoric reference and 2) constrain the candidates down to one choice for a given anaphor (Webber, 1979). This section is organized around those two tasks.

2.1 What referents are available

As information about objects, situations, events, etc. is introduced into a discourse, the addressee constructs a mental model of that information. This set of information is called a *Discourse Model* (DM), and the items within it are *Dis-*

course *Entities* (DE's). Each DE represents a particular concept's participation in the discourse. Most previous anaphora resolution algorithms only create DE's for the referents of NPs; but it is well known that this provides insufficient support for all anaphoric pronouns, for example (inspired by (Webber, 1988)):

(A0) Each Fall, penguins migrate to Fiji.

(A1a) *That's* where they wait out the winter.

(A1b) *That's* when it's too cold even for them.

(A1c) *That's* why I'm going there next month.

(A1d) *It* happens just before the eggs hatch.

The pronouns in A1a and A1b depend only on noun phrases in A0 (Fiji and Fall), but A1c refers to all of A0 characterized as a fact, and the pronoun in A1d refers to the migration event. This example demonstrates that the same discourse unit (A0) can be characterized in a variety of ways by the anaphoric reference (a fact and an event in this case). These referents built from constituents other than noun phrases are called *abstract* referents.

Webber (1990) suggests that each discourse unit produce a pseudo-DE that "stands proxy for its propositional content". Proxies can be coerced into different types of referents by applying *referring functions* (Nunberg, 1979). Each referring function recharacterizes the discourse unit in a particular way. To interpret the pronouns in example A, two functions are needed: one that takes the proxy for A0 and creates a fact referent, and one for the event.¹ Asher (1993) details the syntactic and semantic conditions in which abstract referents are available for anaphoric reference.

In dialog, Walker and Whittaker (1990) point out that once a question is asked, it remains salient until answered. As a result, in problem-solving conversations, particular aspects of the solution that are important for determining how well the proposed solution addresses the overall goal are salient even if they were not explicitly mentioned. These referents must be available

¹Webber (1990) additionally describes a model of discourse structure that explains which discourse units are in focus at any point in the discourse, but PHORA does not employ a sophisticated model of discourse structure so those details are outside the scope of this discussion.

for interpreting anaphoric pronouns. Which aspects of the problem state are salient after each action depends on the task underway. These salient aspects of the solution can be inferred using referring functions in a process similar to evoking abstract referents. For example, in the TRAINS93 domain, when the discourse participants must determine how much time the plan will take, the following exchange is typical:

(B0) Send engine 1 to Elmira.

(B1a) That's six hours. / That'll be six p.m.

DE's evoked by NP's remain in the DM for the duration of a discourse to support long-distance anaphora. Abstract entities are only available for pronominal reference in the immediately following discourse unit (Passonneau, 1991).

2.2 Pronoun Resolution Techniques

Most extant algorithms operate by linking an anaphoric pronoun with the noun phrase that is judged to be its antecedent. The search is ordered either using syntactic structure (Hobbs, 1986) or an estimation of discourse salience (*e.g.* (Brennan et al., 1987; Baldwin, 1997; Tetreault, 2001)). Although different techniques use different evidence for judging salience (grammatical role, information structure, coherence structure, etc.), it is generally agreed that items that appear as NP's in the sentence are most salient.² Other items, such as the situation described by the sentence, are less salient. Therefore, pronouns that refer to these less salient items are simply not addressed in the salience paradigm because they cannot be resolved correctly.

The salience paradigm was originally motivated by the fact that some pronouns have more than one candidate referent with equally valid semantic properties, yet human readers do not find the sentence ambiguous. In example C from Sidner (1983, pg. 373), the pronoun in sentence C1 can refer to either the dog or the bull, but the dog is chosen because it is more focused. Sidner explains that the focused entity is chosen as the referent for a pronoun unless its semantic properties are incompatible with the predicative

²The reader is referred to (Tetreault, 2001) for a survey and evaluation of several salience-based algorithms.

context of the pronoun. When multiple entities with equal semantics are available as candidate antecedents, salience provides a useful heuristic for choosing between them.

(C0) Pam walked her *dog* near *a bull* one day.

(C1) *He* trotted quietly along.

Relying on salience to guide pronoun interpretation does not work for demonstrative pronouns, which differ from personal pronouns along several interrelated dimensions. First, they tend to have clausal antecedents or NP antecedents that were not in subject position (Pasonneau, 1993). Second, they tend to refer to semantically complex items for which the speaker cannot quickly compute agreement features that would be required for a personal pronoun (Channon, 1980). Notice the subtle difference between D0a and D0b. The demonstrative pronoun in D0b refers to the complex object apple+napkin. Is that object plural or singular? The demonstrative pronoun allows the speaker to hedge:

(D0a) Put the apple on the napkin and then move *it* to the side.

(D0b) Put the apple on the napkin and then move *that* to the side.

Third, demonstrative pronouns tend to refer to an entity not in focus (Linde, 1979). Schuster (1988) found that her subjects interpreted sentences like E1a and E1b differently:

(E0) John thought about {becoming a bum}_{*i*}.

(E1a) *It*_{*i*} would hurt his mother and *it*_{*i*} would make his father furious.

(E1b) {*It*_{*i*} would hurt his mother}_{*j*} and *that*_{*j*} would make his father furious.

The first *it* in *E1a* refers to “becoming a bum” and thereby focuses attention on this referent. A second *it* maintains that reference, while the demonstrative pronoun prefers an unfocused entity. This contrast was formalized in the Givenness Hierarchy developed by Gundel et al. (1993), which distinguishes the *in focus* status, marked by personal pronouns, from the *activated* status, marked by demonstratives. Activated entities are in short-term memory because they were mentioned in the discourse or because they are in the physical context, but are not currently at the center of attention.

Because of these differences, anaphoric demonstrative pronouns have typically been excluded from pronoun resolution studies. A notable exception is the model by Eckert and Strube (2000), which matches anaphoric pronouns with their antecedents (pronouns for which no antecedent is identified are labeled *vague* pronouns). This algorithm uses the predication context of the pronoun to label it as incompatible with either abstract (clausal) or individual (noun phrase) referents. For example, the pronoun in “That’s right” would be labeled individual-incompatible. When the predication does not constrain the pronoun, demonstratives prefer abstract referents and personal pronouns prefer individuals.

Their pronoun resolution technique is an extension of the S-list technique (Strube, 1998), which incrementally updates the relative salience of NP-evoked DEs based on information structure, and judges the salience of clausal antecedents using discourse structure. This model was designed for spoken dialog, so it requires antecedents to be in the discourse common ground before they are available for anaphora resolution. This model assumes that referring functions are available to coerce clausal antecedents into the intended referents, but does not specify the details for those functions. Although not implemented, this technique performed well in hand-simulation on a set of Switchboard dialogs.

3 Details of PHORA

PHORA can be summarized as three steps, each of which is detailed in this section:

1. Build DE proxies for Discourse Unit (DU)_{*n*}.

To resolve each pronoun in DU_{*n+1*}:

2. Calculate the most general semantic type T that satisfies constraints on the predicate argument positions the pronoun appears in.
3. Test DE’s in salience order to find a referent that matches the pronoun’s agreement features, using different search orders for personal and demonstrative pronouns. Testing each DE involves applying the referring

Table 1: Constituents that evoke activated entities (DE proxies)

Trigger Constituent	Example
Infinitive or gerund phrase	To load the boxcars/Loading them takes an hour.
The entire clause	I think that he’s an alien.
That sentential	I think that he’s an alien.
Subordinate clause	If he’s an alien...

functions in an attempt to find a referent of type T or one of its subtypes.

3.1 Incrementing the Discourse Context

The DM is partitioned into two sets: Mentioned Entities and Activated Entities. Mentioned DEs carry the interpretation of noun phrases and are evoked as soon as the NP is interpreted. After interpreting each DU, which in this model is each clause, one Mentioned DE is designated as the *focus*.³ Other constituents, such as entire sentences and nominals, create DE proxies in the Activated Entities list. DE’s that represent referents with no existential presupposition (indefinites, commands, etc.) are not available as referents until the utterance evoking them is Confirmed. DE’s have these attributes:

- Input: The surface linguistic constituent
- Number: singular or plural
- Type: such as ENGINE, PERSON, etc.
- Composition: hetero- or homogeneous⁴
- Specificity: individual or kind
- Interpretation: the referent or Proxy associated with this DE

Mentioned entities All referential NP’s⁵ create DE’s with mentioned salience. This includes proper names, descriptive NP’s, and demonstrative and 3rd-person personal, possessive, and reflexive pronouns.

³This model does not commit to a particular method of calculating the focus. For the evaluation reported in Section 4, the left-most NP in each clause was designated as the focus.

⁴A plural entity has heterogeneous semantics if its elements are of different semantic types, such as a set of engines, boxcars, and tankers.

⁵Constituents that function syntactically as NP’s, but that do not refer, do not evoke mentioned entities. Examples are predicate complements, frequency adverbials and expletives (as in “It’s good that you cleaned up”).

Activated entities Activated DEs, containing proxies for linguistic constituents other than noun phrases, are evoked after interpreting each clause. Some example rules for creating proxies are shown in Table 1. Notice that one clause might trigger multiple proxies. During pronoun interpretation, a set of referring functions can be invoked to coerce the proxy into a referent of the desired type. Table 2 lists some example referring functions. These functions are sensitive to both the syntactic and semantic analysis of the original constituent, as well as its speech act. The speech acts defined in this model are:

Tell: Assert the truth of a proposition.

Request: Command one’s partner to perform an action.

WH-Question: Elicit information from one’s discourse partner.

YN-Question: Elicit confirmation of an assertion from one’s discourse partner.

Confirm: Accept an utterance made by one’s partner.

Speech acts are included because they determine which abstract entities are available for subsequent reference. For example, WH-questions cannot be characterized as propositions by subsequent anaphoric reference:

(F0a) The highway is closed (Tell)

(F0b) Is the highway closed? (Y/N Question)

(F1) That’s right.

(F0b) Why is the highway closed? (WH-Q)

(F1) *That’s right.

After each clause that updates the solution being constructed (TELLs and REQUESTs), a proxy DE is added for the plan. Reference functions are defined to infer salient task-related entities such as the completion time of the plan. For instance, if a sentence describes an event

Table 2: Referring Functions

Function	DE / Proxy Details	Output Types
Ident(d,t)	Any Mentioned DE if its type meets type constraint t	The train is red.
Kind(d,t)	Descriptive NP (not bare plural NP) meeting of type t	That's a great route .
Situation(d)	Sentence with tensed stative verb	The train is red.
Event(d)	Sentence with tensed eventive verb	It gets there late.
Kind _A (d)/Kind _E (d)	infinitive form of action/event	To load them takes an hour.
	Gerund form of action/event	Loading them takes an hour.
Proposition(d)	Each TELL or YN-Question that sentential	I think that he's an alien.
	if/when subordinate clause	I think that he's an alien .
Endtime(d)	Sentence with tensed eventive verb	If he's an alien...
	Gerund or Infinitive from eventive verb	Then we go to Avon.
		I need to load the boxcars.

to be added to the plan, and the task must be completed in a certain amount of time, the completion time of that task is salient. The function Endtime(d) would return a referent of type TIME-POINT denoting the end time of the event for which d is a proxy.

Mentioned DE's remain in the DM for the entire discourse. Activated DE's from Clause_n remain only while interpreting Clause_{n+1}. The anaphoric mention of an abstract referent evokes a Mentioned DE that is treated just like any other Mentioned DE.

3.2 Semantic Constraints

Semantic type constraints for each pronoun are determined as the first step of pronoun resolution. In some contexts, the pronoun's type is left unconstrained. For example, in "That's good" the pronoun is the argument of ACCEPTABLE(X), which can apply to any semantic type. Semantic constraints come from:

Verb senses have associated semantic restrictions for each argument position. (e.g. "Load them into the boxcar." produces the constraint that the theme of LOAD must be CARGO).

Predicate NPs Be-verbs and other near-copular constructions that are interpreted as EQUAL constrain the subject to be the same type as the complement. For example: "That's a good route" constrains the subject to be a ROUTE.

Predicate Adjectives Copulas containing adjectives also constrain the possible se-

mantic type of a pronoun in subject position. For example, "it's right" is interpreted as CORRECT(X) and the argument of CORRECT must be a PROPOSITION.

3.3 Pronoun resolution search order

Pronoun resolution is a search for the first referent that satisfies semantic type constraints and agreement features for the pronoun. A referent satisfies the type constraint if it is of type T or a subtype of that type. The search through Mentioned entities works backwards in the order the clauses appeared in the discourse.⁶ Within each clause, Mentioned entities are searched in left-to-right order.⁷ Searching the Activated entities means applying referring functions that produce an acceptable type of referent to each DE proxy. The referring functions produce a referent if the function was successful or NIL if the function was not able to produce a referent of the desired type from the proxy DE. Personal and demonstrative pronouns employ different search orders.

Personal The referent of a **singular pronoun** is the first singular, individual referent found in this order:

- Mentioned entities to the left of the pronoun in the current clause are searched in right-to-left order
- The focused entity from Clause_{n-1}

⁶Although linear order is an inaccurate model of discourse structure (Grosz and Sidner, 1986; Webber, 1988; Walker, 1996), no method exists for building an accurate discourse structure in realtime.

⁷Using the breadth-first ordering of Hobbs algorithm instead is a point for future work.

Table 3: DM after “It takes an hour to load the oranges.”

Input	Number	Type	Composition	Specificity	Interpretation	Salience
Engine 1	Sing	ENGINE	Homogeneous	Indiv	ENG1	focus
Avon	Sing	CITY	Homogeneous	Indiv	AVON	mentioned
the oranges	Plural	ORANGE	Homogeneous	Indiv	ORANGES1	mentioned
to get oranges	Sing	Functional	Homogeneous	Kind	proxy	activated
All of G0	Sing	Functional	Homogeneous	Indiv	proxy	activated
The plan	Sing	Functional	Homogeneous	Indiv	proxy	activated

- The remaining Mentioned entities
- Activated entities from Clause_{*n*-1}

For **plural pronouns**, the algorithm uses the same search order but looks for plural referents. If no plural entity is found, it applies the Kind(d) referring function to each Mentioned DE in an attempt to generate a Kind referent.

Demonstrative Pronouns *This/that* may refer to Kinds, singular or plural referents, while *these/those* cannot refer to singular referents. The referent is the first one meeting number agreement using this search order:

- Activated entities from Clause_{*n*-1}
- The focused entity only if it can be coerced to a Kind
- Mentioned DE’s from the entire discourse (preferring heterogeneous ones)

3.4 An Example

This section demonstrates PHORA at work. Table 3 shows the DE’s evoked by G0 that provide possible referents for G1. AVON, ENG1, and ORANGES1 are objects from the referential world. MOVE and LOAD are actions.

(G0) Engine 1 goes to Avon to get the oranges. G0 is interpreted as (this is an abbreviated version of the logical form):

```
(TELL (MOVE :theme x :dest y
        :reason (LOAD :theme w)))
(the x (refers-to x ENG1))
(the y (refers-to y AVON))
(the w (refers-to w ORANGES1))
```

(G1a) So it’ll get there at 3 p.m.

```
(ARRIVE :theme x :dest y :time z)
ARRIVE constrains its theme to be a
```

MOVABLE-OBJECT, so the candidates are {ENG1, ORANGES1}. The search order for *it* finds ENG1 as the referent.

(G1b) That takes two hours.

```
(TAKE-TIME :theme x :cost y)
TAKE-TIME constrains its theme to
EVENTs. The search order for that at-
tempts Activated DEs first, using the func-
tion Event(d). The function is successful on
the proxy for G0 since MOVE is an event.
```

(G1c) That’s where the orange warehouse is.

```
(EQUAL :theme x :complement y)
x must be a LOCATION since that is
the type of the complement where clause.
No function creates locations, so the
search proceeds to Mentioned entities, and
{AVON} satisfies the type constraint.
```

4 Evaluation

PHORA was evaluated on a set of ten problem-solving dialogs from the TRAINS93 corpus (Heeman and Allen, 1995). The meaning of all third-person pronouns in these dialogs had been determined in a previous study (Byron, 2000).⁸ The resulting dataset contains 557 utterances and 180 test pronouns.⁹

PHORA runs as the last step of semantic analysis in the TRIPS spoken dialog system. TRIPS’ grammar and lexicon coverage was extended to parse these dialogs. In addition, minor changes were made to the dialogs to allow them to parse, but no changes were made to details of interest to PHORA such as the phrasing of nominals, the exact form of referring expressions, the

⁸Judging the referent of demonstrative pronouns had a Kappa score of only .56 in that study, therefore the answer key was derived by two annotators agreeing on the correct referent.

⁹No gendered pronouns and no instances of *its*, *itself*, *their*, or *these* appear in these ten dialogs.

Table 4: Pronoun Resolution Success on 10 TRAINS93 dialogs

	It	They	Them	T'selves	Total PRP	That	This	Those	Total DEM	Out of Scope	Total
A Raw word count	89	19	10	1	119	127	2	18	147	326	592
Non-Referential Tokens											
B Expletive	32	0	0	0	32	2	0	0	2	0	34
C Determiners/Relative Pron	0	0	0	0	0	25	1	8	14	0	14
D Sum Non-Referential (B+C)	32	0	0	0	32	27	1	8	36	0	68
E Referential Pronouns (A-D)	57	19	10	1	87	100	1	10	111	0	524
Referential Tokens That were Excluded											
F First/Second person	0	0	0	0	0	0	0	0	0	326	326
G Abandoned Utterance	7	0	0	0	7	7	0	4	11	0	18
H Total Excluded (F+G)	7	0	0	0	7	7	0	4	11	326	344
I Evaluation set (E-H)	50	19	10	1	80	93	1	6	100	0	180
1 LRC Baseline	27	16	8	1	52	10	0	4	14	0	66 (37%)
2 Adding semantic type constraint	32	14	11	1	58	14	0	5	19	0	77 (43%)
3 Adding abstract referents	37	15	11	1	64	51	0	5	56	0	120 (67%)
4 Demonstr. different search order	37	15	10	1	63	62	0	5	67	0	130 (72%)

verb commanding a test pronoun, etc. The semantic resources needed for the dialogs (i.e. entries in the semantic type hierarchy, type constraints on predicate argument positions) were also built for this evaluation. Each pronoun was considered correctly resolved if both the correct antecedent and the correct referring function were chosen. All pronoun interpretation and DM construction was performed automatically by the PHORA software.

4.1 Results

There is no other pronoun resolution software that resolves non-NPC pronouns to use for comparison in this evaluation. Eckert and Strube’s model relies on hand-annotated discourse structure, which was not available for this corpus. Therefore, the LRC algorithm, which performs well on NPC pronouns (Tetreault, 2001), was chosen as the baseline. LRC uses salience only to select NP antecedents, and treats demonstratives the same as personal pronouns. Rows A through I of Table 4 show the composition of the evaluation corpus using the format defined in (Byron, 2001). Pronoun resolution results start with LRC in the row numbered 1, and the different elements of PHORA are gradually added to this baseline technique. LRC correctly resolved 37% of pronouns in the evaluation set.¹⁰

¹⁰Only 50% of pronouns in this set are NPC, so that is the upper limit for LRC on this data. LRC resolves 69% of the NPC pronouns in this test set correctly, which is

Row 2 shows the improvement gained by adding semantic type matching, which improves both personal and demonstrative pronouns. Row 3 shows the effect of adding activated entities for non-NP constituents. The success rate on demonstratives is expected to improve, but notice that personals improve by 7% as well. Finally, the full PHORA model is shown in Row 4. Using a different search order for demonstrative pronouns correctly resolves 72%, nearly doubling the LRC baseline. Many of the remaining errors were due to the simplistic linear model of discourse structure. Detecting embedded dialogs or utilizing prosody to detect structural shifts (Hirschberg and Pierrehumbert, 1986), should improve performance and is proposed as future work.

One limitation of this model is its dependence on semantic resources. A domain-independent evaluation was conducted by labeling semantic constraints as domain-dependent or independent (e.g. the theme of “that’s right” is a proposition in any domain, but the theme of “they go to Avon” as a MOVABLE-OBJECT is specific to TRAINS93). On the same evaluation data with only domain-independent semantics enabled, the full PHORA model resolved 51% of the test pronouns correctly. Applying the model to open-domain discourse using a standard semantic resource, such as Wordnet, is an anticipated future direction.

comparable with its performance on other datasets.

5 Summary

The pronoun resolution technique described here significantly expands the range of anaphoric pronouns that can be automatically interpreted. It draws upon many previous lines of research to allow abstract entities to be generated as candidate referents for anaphoric pronouns, and adds semantic filtering during pronoun resolution to counteract this increase in potential referents. To date, the model has been implemented only in a closed-domain system. However, the technique itself is not domain-specific, and testing the technique on open-domain discourse is a point of future work.

Acknowledgements

I would like to thank James Allen, Len Schubert, Greg Carlson, and Candy Sidner for advice and support on this project. This material is based on work supported by ONR grant N00014-01-1-1015; DARPA grant F30602-98-2-0133; and NSF grant EIA-0080124. Opinions, findings and recommendations expressed here are those of the author and not the views of these fine organizations.

References

- N. Asher. 1993. *Reference to Abstract Objects in Discourse*. Kluwer Academic Publishers.
- B. Baldwin. 1997. Cogniac: High precision coreference with limited knowledge and linguistic resources. In *Operational Factors in Practical, Robust Anaphora Resolution for Unrestricted Texts (ACL-97 workshop)*, pages 38–45.
- S. Botley. 1996. Comparing demonstrative features in three written english genres. In *Proceedings of the Discourse Anaphora and Resolution Colloquium (DAARC96)*, pages 86–105.
- S. Brennan, M. Friedman, and C. Pollard. 1987. A centering approach to pronouns. In *Proceedings of ACL '87*, pages 155–162.
- D. Byron and J. Allen. 1998. Resolving demonstrative pronouns in the TRAINS93 corpus. In *Proceedings of the Second Colloquium on Discourse Anaphora and Anaphor Resolution (DAARC2)*, pages 68 – 81.
- D. Byron. 2000. Analysis of pronominal reference in two spoken language collections. Technical Report 703, University of Rochester Computer Science Dept.
- D. Byron. 2001. The uncommon denominator: A proposal for consistent reporting of pronoun resolution results. *Computational Linguistics*, 27(4).
- R. Channon. 1980. Anaphoric *that*: A friend in need. In J. Kreiman and A. Ojeda, editors, *Papers from the Parasession on Pronouns and Anaphora*, pages 98–109. Chicago Linguistic Society.
- M. Eckert and M. Strube. 2000. Dialogue acts, synchronising units and anaphora resolution. *Journal of Semantics*, 17(1):51–89.
- B. Grosz and C. Sidner. 1986. Attention, intentions, and the structure of discourse. *Computational Linguistics*, 12(3):175–204.
- J. Gundel, N. Hedberg, and R. Zacharski. 1993. Cognitive status and the form of referring expressions in discourse. *Language*, 69(2):274–307.
- P. Heeman and J. Allen. 1995. The Trains spoken dialog corpus. CD-ROM, Linguistics Data Consortium.
- J. Hirschberg and J. Pierrehumbert. 1986. The intonational structuring of discourse. In *Proceedings of ACL '86*, pages 136–144.
- J. Hobbs. 1986. Resolving pronoun reference. In *Readings in Natural Language Processing*, pages 339–352. Morgan Kaufmann.
- C. Linde. 1979. Focus of attention and the choice of pronouns in discourse. In Talmy Givon, editor, *Syntax and Semantics 12: Discourse and Syntax*, New York. Academic Press.
- G. Nunberg. 1979. The non-uniqueness of semantic solutions. *Linguistics and Philosophy*, 3(2):142–184.
- R. Passonneau. 1991. Some facts about centers, indexicals, and demonstratives. In *Proceedings of ACL '91*, pages 63–70.
- R. Passonneau. 1993. Getting and keeping the center of attention. In M. Bates and R. Weischedel, editors, *Challenges in natural language processing*, pages 179–226. Cambridge University Press.
- E. Schuster. 1988. Anaphoric reference to events and actions: Evidence from naturally-occurring data. Technical Report MS-CIS-88-13, University of Pennsylvania LINC LAB.
- C. Sidner. 1983. Focusing in the comprehension of definite anaphora. In M. Brady and R. Berwick, editors, *Computational Models of Discourse*, pages 363–394.
- M. Strube. 1998. Never look back: An alternative to centering. In *Proceedings of ACL '98*, pages 1251–1257.
- J. Tetreault. 2001. A corpus-based evaluation of centering and pronoun resolution. *Computational Linguistics*, 27(4):507–520.
- M. Walker and S. Whittaker. 1990. Mixed initiative in dialogue: An investigation into discourse segmentation. In *Proceedings of ACL '90*, pages 70–78.
- M. Walker. 1996. Limited attention and discourse structure. *Computational Linguistics*, 22(2):255–264.
- B. Webber. 1979. *A Formal Approach to Discourse Anaphora*. Garland Publishing.
- B. Webber. 1988. Discourse deixis: reference to discourse segments. In *Proceedings of ACL '88*, pages 113–122.
- B. Webber. 1990. Structure and ostension in the interpretation of discourse deixis. Technical Report MS-CIS-90-58, Department of Computer and Information Science, University of Pennsylvania.