# Associative Descriptions and Salience: A Preliminary Investigation

**Massimo Poesio**
Department of Computer Science, University of Essex
poesio@essex.ac.uk

## Abstract

We discuss the problems involved in identifying and annotating bridging descriptions in corpora of English, and present results concerning the correlation between bridging descriptions and Centering obtained using a reliably annotated corpus and automatic focus-tracking methods.

## 1 Introduction and Motivations

In previous work (Poesio et al., 1997; Poesio and Vieira, 1998; Poesio et al., 1998; Vieira and Poesio, 2000; Poesio et al., 2002a) we studied the use of BRIDGING REFERENCES (Clark, 1977) and other types of definite descriptions in corpora of English, and developed methods for identifying their ANCHOR[1] exclusively based on lexical information and simple window-based segmentation methods. Three questions for future research were identified in that work:

1. When the full range of bridging references identified by Clark is considered, subjects have serious difficulty agreeing on bridging references, and are only able to identify very few (Poesio and Vieira, 1998);

2. Even large lexical resources such as WordNet (Fellbaum, 1998) do not contain all the necessary information for resolving Bridging Descriptions (BDs) even if only those BDs that rely on the type of relations found in WordNet such as hyponymy, synonymy, or meronymy (e.g., *the window* / a *house*) are considered. In almost 40% of the cases, no semantic relation could be found between such bridging references and their anchor (Poesio et al., 1997; Vieira and Poesio, 2000);

3. Even when a lexical resource did contain information about the existence of a relation between a bridging description and its anchor (30% of the total with WordNet), simply choosing as anchor the entity which is semantically closer to the BD in the chosen window leads to very poor results. In about half of the cases, a competing discourse entity was semantically closer than the actual anchor. We concluded that the resolution of a bridging description does not depend only on lexical information, but also involves other information–perhaps, about which entities are most salient, either in the basic sense of being more recent, or perhaps by being the 'focus', as already claimed by Sidner (1979).

We addressed the second of these problems by developing methods for automatically acquiring the lexical information needed to resolve (a subset of) BDs (Poesio et al., 1998; Poesio et al., 2002a). In parallel with that effort, we have been developing reliable methods for annotating bridging references, and created a corpus which can be be used to study them with greater accuracy (Poesio, 2000). We have also been developing automatic methods for tracking the 'focus' of a discourse in the sense of Centering Theory (Grosz et al., 1995; Walker et al., 1998) which allow us to compare the many existing definitions of notions such as 'utterance', 'ranking,' 'Backward-Looking Center,' etc., identifying those that resulted in fewer violations of the theory's claims

[1] We borrow the term anchor from Fraurud (1990) to indicate the discourse entity with which the bridging reference is 'associated' - reserving 'antecedent' for identity anaphora.

(Poesio et al., 2000; Poesio et al., 2002b). In the work reported here, we exploited our corpus and these focus-tracking techniques to study the correlation between 'salience' and bridging reference resolution, focusing on ASSOCIATIVE DESCRIPTIONS (ADs)–i.e., bridging references realized as *the*-NPs, related to their anchor by a relation other than identity (Hawkins, 1978). The structure of this paper is as follows. In section 2, we discuss the corpus used in this study, how it was annotated, and the methods used to track the CB in this corpus. In section 3, we present our results concerning the correlation between bridging references and salience. Finally, in section 4, we present preliminary findings concerning the use of lexical semantics as a filter.

## 2 Methods

One of the main motivations for this work is that we felt that we could improve upon our previous results by exploiting the results of our other work on referring expressions in general and of salience (Poesio et al., 2000; Poesio, 2000). As a result of this work we had at our disposal the GNOME corpus whose NPs, and the anaphoric relations between them, had been marked in a reliable way (Poesio, 2000). In this section we briefly discuss the corpus and how it was annotated.

### 2.1 The Corpus

The GNOME corpus (Poesio, 2000) includes texts from three domains, two of which were used in this study. The museum subcorpus consists of descriptions of museum objects and brief texts about the artists that produced them.[2] The pharmaceutical subcorpus is a selection of leaflets providing the patients with mandatory information about their medicine.[3] Each subcorpus contains about 6,000 NPs; in this study we used texts from the first two domains, for a total of about 3,000 NPs, about 600 of which are definite descriptions. Of the potential candidates for 'utterances', the corpus includes about 500 sentences, and 900 finite clauses; the actual number of utterances used in the study is one of the parameters that we varied.

[2]The museum subcorpus extends the corpus collected to support the ILEX and SOLE projects at the University of Edinburgh. ILEX generates Web pages describing museum objects (Oberlander et al., 1998) The SOLE project extended ILEX with concept-to-speech abilities (Hitzeman et al., 1998).

[3]The leaflets in the pharmaceutical subcorpus are a subset of the collection of all patient leaflets in the UK , digitized to support the ICONOCLAST project at the University of Brighton (Scott et al., 1998).

### 2.2 Annotation

**Motivations** The marking scheme used in (Poesio and Vieira, 1998) wasn't completely satisfactory. For one thing, our method for annotating bridging references didn't guarantee agreement. Secondly, although our work had revealed that many definite NPs could simultaneously belong to two classes - e.g., be directly anaphoric on one entity, while bridging on another one–our scheme wouldn't allow our annotators to mark this information. Finally, we didn't have a way for marking genuinely ambiguous cases. These problems were avoided by the marking schemes developed for the GNOME and MATE projects. The following example from the GNOME corpus illustrates how anaphoric expressions may be related to more than one anchor in sometimes complex ways:

(1) a. *Each coffer* also has *a lid* that opens in two sections.
b. The upper lid reveals a shallow compartment,
c. while the main lid lifts to reveal the interior of the coffer.
d. The 1689 inventory of the Grand Dauphin, the oldest son of Louis XIV, lists *a jewel coffer of similar form and decoration*;
e. according to this inventory, André-Charles Boulle made the coffer.
f. The two stands are of the same date as the coffers, but were originally designed to hold rectangular cabinets.

Notice, first of all, that both *the upper lid* in (1b) and *the main lid* in (1c) are related to both *each coffer* and *a lid* in (1a). Determining the exact relation is made more difficult by the fact that these are examples of so-called *telescoping*–when an anaphoric expression appears to be 'in the scope' of a quantified expression even though that scope should have been 'closed off' by the end of the sentence–but the relation with the coffer is almost

certainly a PART-OF relation, whereas the relation with the lid as a whole may be viewed either as PART-OF or perhaps SET-ELEMENT. Next, notice *the coffers*

**Basics** The GNOME corpus was annotated according to a systematic manual, available from the GNOME project's home page at http://www.hcrc.ed.ac.uk/~gnome. The annotation scheme derives from the MATE scheme (Poesio et al., 1999). Only the most important details of the scheme will be discussed.

All sentences in the GNOME corpus are marked as elements of type ⟨s⟩. In addition, all units of text in the corpus that might be identified with utterances (in the Centering sense) are marked as ⟨unit⟩ elements. Each NP is marked with a ⟨ne⟩ tag and with a variety of attributes capturing syntactic and semantic properties.

**Marking Bridging References and other Anaphors** The problem with bridging references is that many NPs evoke discourse entities that are related to other discourse entities, without truly establishing a 'bridge' with the rest of the discourse. An obvious example of this are NPs indicating the possessor in a possessive NP, such as the definite description *the Sun King* in *the Sun King's possessions*. Clearly, the discourse entity denoted by *the Sun King* is semantically related to the discourse entity denoted by the possessive NP; yet it's also clear that this NP does not establish a link to previous discourse. (Furthermore, identifying the anchor of this NP is not a problem.) In order to be consistent about what we classified as 'bridging reference', we developed, first of all, methods for marking both identity and associative semantic relations; secondly, and simplifying matters somewhat, we only classified as bridging references the NPs related by (non-identity) semantic relations with discourse entities last mentioned in the previous utterance (in the sense of Centering, see below).

In the MATE / GNOME scheme, anaphoric relations are not marked using attributes of the ⟨ne⟩ element, as in the MUC scheme and most other schemes, but using a separate ⟨ante⟩ element. The ⟨ante⟩ element itself specifies the index of the anaphoric expression and the type of semantic relation (e.g., identity), whereas one or more embedded ⟨anchor⟩ elements indicate possible antecedents (the presence of more than one ⟨anchor⟩ element indicates that the anaphoric expression is ambiguous). (See (2).)

(2)
```
<unit finite='finite-yes'>
 <ne id='ne546' gf='subj'>
   The drawing of
   <ne id='ne547' gf='np-compl'>
   the corner cupboard, </ne></ne>
 <unit finite='no-finite'>
  or more probably
   <ne id='ne548' gf='no-gf'>
      an engraving of
      <ne id='ne549' gf='np-compl'>
      it
      </ne></ne>
 </unit>,
...
</unit>
<ante current="ne549" rel="ident">
<anchor ID="ne547">
</ante>
```

Separating the ⟨ante⟩ element from the ⟨ne⟩ element makes it possible to mark an NP both as coreferring with one entity, and bridging on a second one, if necessary: this can be done by specifying two separate ⟨ante⟩ elements. It is also possible to mark ambiguous anaphoric expressions, by specifying more than one ⟨anchor⟩ element.

Previous work, particularly in the context of the MUC initiative, suggested that while it's fairly easy to achieve agreement on identity relations, marking up bridging references is quite hard. This was confirmed In our own previous work, in which we found that our subjects only agreed on about 5% of bridging descriptions (Poesio and Vieira, 1998). We addressed this problem by limiting the types of relations annotators are supposed to mark up, and by specifying priorities. Our annotators only annotated four types of relations, a subset of those proposed in the 'extended relations' version of the MATE scheme (Poesio et al., 1999): identity (IDENT), set membership (ELEMENT), subset (SUBSET), and 'generalized possession' (POSS), which also includes part-of relations. In addition, we asked our annotators to use a special tag OTHER to mark up semantic relations between an anaphoric expression and an entity introduced by an NP in a previous clause or layout element, when such relation existed and no other semantic relation between the ⟨unit⟩ in which the NP occurred

and previous units had been identified.

As expected, we obtained a rather good agreement on identity relations. In our most recent analysis (two annotators looking at the anaphoric relations between 200 NPs) we observed no real disagreements; 79.4% of these relations were marked up by both annotators; 12.8% by only one of them; and in 7.7% of the cases, one of the annotators marked up a closer antecedent than the other. For bridging references, only 4.8% relations were now marked differently; on the other hand, only 22% of bridging references were marked in the same way by both annotators. In other words, although our current scheme does limit the disagreements on antecedents and relations, we still find that 73.17% of relations are marked by only one annotator.

### 2.3 Automatic tracking of focus according to Centering theory

The notion of 'topic' or 'discourse focus' has often been claimed to play an important role in the interpretation and production of anaphoric expressions, including bridging references (Grosz, 1977; Sidner, 1979; Sanford and Garrod, 1981; McKeown, 1986; Grosz et al., 1995; Dale, 1992; Walker et al., 1998), but it is notoriously difficult to pin down. Sidner (1979) developed detailed algorithms implementing her claim (and Grosz's (Grosz, 1977)) that the interpretation of bridging references depend on focusing information. Her proposals however rely heavily on commonsense knowledge, and for this reason have never been extensively evaluated.[4]

As the basis for the work reported here, we used the terminology and ideas introduced in what is currently the best-known formalization of the idea of topic in Computational Linguistics, Centering Theory (Grosz et al., 1995; Walker et al., 1998), in particular the notions of Backward-Looking Center (CB) and Preferred Center (CP). In the 'mainstream' version of Centering (Grosz et al., 1995), it is assumed that each UTTERANCE introduces new discourse entities (or Forward-Looking Centers) into the discourse, and in so doing, updates the 'local focus'. It is further assumed that the discourse entities introduced (or better, REALIZED) by an utterance are ranked; the most highly ranked entity in an utterance is called the CP. The CB is Centering's equivalent of the notion of 'topic' or 'focus', and is defined as follows:

**CB** CB($U_i$), the BACKWARD-LOOKING CENTER of utterance $U_i$, is the highest ranked element of CF($U_{i-1}$) that is realized in $U_i$.

Centering provides no definition of the notions of 'ranking', 'utterance' and 'realization'; researchers using the theory have to specify their own. In previous work, we did a comparative analysis of the many existing proposals concerning Centering Theory's parameters (Poesio et al., 2000; Poesio et al., 2002b). The GNOME corpus, annotated as discussed above, was used to automatically compute 'utterances' according to different definitions proposed in the literature, and then to compute the CFs and the CB (if any) of each utterance on the basis of the anaphoric information and according to different views of ranking. This information was the used to find violations of the main claims of the theory.

One of our results was that two ways of specifying the parameters of the theory gave the best (and pretty much equivalent) results, in the sense that it led to the fewest violations of the main claims of the theory. Both of these involved identifying utterances with sentences, and allowing for indirect realization of the CB; but they differed in the ranking function. In one case, grammatical function was used (subjects rank more highly than objects that rank more highly than adjuncts) augmented with a linear disambiguation factor; in the other, Strube and Hahn's (1999) ranking function based on 'information status,' according to which hearer-old entities are ranked more highly than inferrables, which in turn are ranked more highly than hearer-new entities (Prince, 1992).

## 3 Salience and Bridging Reference

The annotated corpus and the focus tracking scripts discussed above were used to investigate the extent to which the interpretation of BDs depended on the salience of potential anchors. In this section we discuss our main findings.

[4] Carter (1987) actually implemented a system with the necessary commonsense knowledge to test Sidner's theory, but given that all the necessary knowledge had to be provided by hand, the system was only tested on a restricted –although not small–number of examples.

### 3.1 Bridging descriptions in the GNOME corpus

A total of 2073 semantic relations are marked up in the GNOME corpus. These can be classified as follows, where IDENT is identity, and POSS-NPs are the relations between the possessor and the possessee in a possessive NP:

| Type of relation | Number |
|---|---|
| IDENT | 1164 |
| POSS-NPs | 328 |
| SUBSET, ELEMENT, POSS, OTHER | 581 |
| Total | 2073 |

Several types of NPs can be used to introduce objects related to previous discourse entities by the type of 'associative' relations usually involved in bridging reference: definite descriptions, possessive NPs, demonstratives, and even quantifiers. Of the 581 'other' relations in the previous table, 194 are expressed by definite descriptions; 169 definite descriptions in total enter into such relations.[5] If we identify utterances with sentences, 110 of these associative description have their anchor in utterances preceding the current one (henceforth, U-1), and could thus be classified as 'bridging' under the simplified definition used here. If utterances are identified with finite clauses, we have instead 128 BDs. The statistics concerning the distance between an associative description and its closest anchor, with sentences as utterances, can be summarized as follows:

| Distance | Number of BDs whose anchor is that far |
|---|---|
| 0 (same utterance) | 59 |
| 1 | 72 |
| 2 | 20 |
| 3-5 (total) | 11 |
| 6 or more | 7 |

(Notice that 95.8% of the anchors are found in the same utterance or the previous 5, confirming the results of (Poesio and Vieira, 1998).)

### 3.2 Bridging descriptions and recency

The simplest hypothesis concerning BD resolution is that their anchors are identified searching first in the previous utterance, then in the utterance before that, etc.–i.e., following the search strategy first proposed in (Winograd, 1972; Hobbs, 1978; Sidner, 1979). Utterances may be searched left-to-right or right-to-left ('most recent first').

In our corpus, 68% (49 / 72) of associative descriptions whose anchor is in the previous sentence have the first-mentioned entity (henceforth, FM(S-1)) as their anchor, and 44.5% of the total. Of the ADs whose anchor is in the current or previous *clause*, 64% have as anchor the first-mentioned element of the previous clause. By constrast, for only 15% (11 / 72) of BDs whose anchor is in S-1, this anchor is the last-mentioned entity in S-1. These figures confirm the preference for first-mention over recency often observed in the literature (Gernsbacher and Hargreaves, 1988; Gordon et al., 1993). In fact, even entities in second and third position are significantly more preferred than those in last position (23.6%, 17 / 72). These figures are summarized by the following table:

| Position of anchor | Number of ADs whose anchor is in that position |
|---|---|
| 1 | 49 |
| 2-3 | 17 |
| last | 11 |
| Total | 72 |

In other words, if you have perfect knowledge about which definite description is bridging, but don't know anything else, choosing as anchor of a BD FM(U-1) will give you 44.5% accuracy. This may not seem much, but it's better than the results obtained with either of the strategies of choosing the anchor on the basis of lexical closeness that we tried in the past, looking for the closest sense in WordNet, which resulted in a 39% accuracy (Poesio et al., 1997), and looking for the closest element in a vectorial space (Poesio et al., 1998), with an accuracy of 22.2%. To do better than that, additional information is needed to evaluate the suggestions made by the search strategy.

### 3.3 Bridging descriptions and focusing

A more complex hypothesis concerning the search for the anchor of a BD is that it involves information about which entities are in focus (Sidner, 1979). Using Centering terminology, and ignoring direct anaphors and more complex cases, Sidner's algorithm (more precisely, her 'Associated Specification', 'Inferred Specification,' and 'Set-Element Specification' rules) could be reformulated as follows ((Sidner, 1979), p. 123-124):

[5] As discussed above, an anaphoric expression may be semantically related to more than one discourse entity.

1. Try first as anchor CB(U-1) (or CP(U-1));
2. If that fails, try all other CFs in ranking order;
3. If that's not acceptable, try the stacked foci.

If we identify utterances with sentences and we use Strube and Hahn's ranking function, then CB(U-1) is the anchor of 37 out of the 72 BDs whose anchor is in the previous utterance (51.3%), and 33.6% overall (37 / 110). Virtually identical results are obtained if grammatical function (+ a linear disambiguation function) is used for ranking. The results are slightly better when the CP is considered, and with Strube-Hahn ranking: 38.2% of BDs have CP(U-1) as their anchor (58% of those whose anchor is in U-1). (With grammatical function ranking, only 22 BDs have CP(U-1) as their anchor.) The results are also unchanged if finite clauses are used instead of sentences: 40 out of 119 BDs, 33.6%, have CB(U-1) as their anchor, and 42 out of 119 have CP(U-1). Clearly, simply choosing the CB (or the CP) as the anchor doesn't work very well, although choosing the CP would give slightly better results. These results are summarized (for sentences) by the following table:

| Focusing status | Number ADs | Overall perc. | Perc. S-1 |
|---|---|---|---|
| CB(S-1) | 37 | 33.6% | 51.3% |
| CP(S-1) | 42 | 38.2% | 58% |
| CB(S') or CP(S'), for some S' | 93 | 84.5% | |
| CB(S) | 70 | 63.6% | |
| Total | 110 | 100% | 100% |

Clearly, simply choosing as anchor the CB or CP of the previous sentence or finite clause doesn't work very well - not even as well as choosing FM(S-1). An example of AD whose anchor is neither CB(S-1) nor CP(S-1) is (1f): the anchor of the AD *the two stands* is the coffers, whereas the CB and CP of S-1 is the inventory. The following is an example in which focusing information helps: the anchor of the AD *the central door* is *this monumental cabinet*, which is both the CB and the CP of (3a) (only with Strube-Hahn ranking):

(3) a. The decoration on *this monumental cabinet* refers to the French king Louis XIV's military victories.
b. A panel of marquetry showing the cockerel of France standing triumphant over both the eagle of the Holy Roman Empire and the lion of Spain and the Spanish Netherlands decorates the central door.

One bit of information coming from focusing, however, appears to be very useful. With sentences as utterances, and either Strube / Hahn or grammatical function ranking, 98 out of 110 anchors of associative descriptions, 89%, are entities that have previously been CB or CP. The results are similar if we identify utterances with finite clauses: in this case, 106 out of 119 anchors, 89% again, were previously CBs (the percentage for CPs is still high, but significantly lower, 82.3%). An example of AD whose anchor is an entity which serves as the CB or CP of a sentence other than S-1 is again *the two stands* in (1f). In fact, close analysis of the 12 cases in which the anchor never was either a CB nor a CP reveals that in virtually all cases the annotation was rather difficult, and there are only two clear cases of AD whose anchor had never before been either a CB or a CP–*the skin* and *the treated area* in (4).

(4) a. Side effects may occur when PRODUCT-X is applied to large areas of *the body*, and for long periods of time (more than 4 weeks), especially if waterproof dressings are used.
b. These include: thinning of the skin on or around the treated area.

Clearly, limiting the search to past CBs, even if following a simpler recency order, would pay off.

To conclude, simply knowing CB(U-1) or CP(U-1) is not even as useful by itself as knowing FM(U-1), but limiting the search to previous CBs might be very useful, provided that ways are found of 'checking' that the proposed solution is plausible, as hypothesized by Sidner.

## 4 Filtering using lexical knowledge: preliminary investigations

Although the results above are not terribly impressive, what Sidner (1979) actually proposed is that anaphors, including bridging references, are resolved by a two-stage process whereby first hypotheses are suggested on the basis of focusing information, and then commonsense inference is used to accept or reject these hypotheses. In this section we briefly consider how easy such a filtering strategy would be to implement.

### 4.1 Using WordNet

Our earlier results suggested that implementing Sidner's filter is not going to be too easy if all we have is WordNet (Poesio et al., 1997; Vieira and Poesio, 2000); preliminary studies with the GNOME corpus confirm this finding. The 110 definite descriptions with an associative interpretation and that are truly 'bridging' to a previous utterance enter into a total of 97 ELEMENT, SUBSET, or POSS relations with an anchor in one of the previous utterances. (Often, with more than one.) We ran a script attempting to find a direct lexical connection in WordNet between any one of the senses of the BD and any one of the senses of the potential anchors in the previous utterances. The script follows the search algorithm discussed above: it starts from the most recent utterance and goes backwards; in each utterance, it considers the CFs left-to-right, following the order found to be best in the tests discussed above; and it stops when it finds a direct link. All of the part-of relations encoded in WordNet (PART_MERONYM, MEMBER_MERONYM, SUBSTANCE_MERONYM, PART_HOLONYM, MEMBER_HOLONYM, and SUBSTANCE_HOLONYM) are considered as possible links. The results are very bad: in only 6 cases the script found a direct lexical connection; and none of these 6 CFs was a plausible anchor for the BD.

Can these results be improved by simply adopting a more extended search strategy? We think that this is unlikely. Consider the following example:

(5) a. This table's marquetry of ivory and horn, painted blue underneath, would have followed *the house's* blue-and-white color scheme, imitating blue-and-white Chinese porcelain, a fashionable and highly prized material.

b. Blue-and-white ceramic tiles decorated *the house*, and some of the furniture was also painted blue-and-white.

The anchor for the BD *the furniture* is clearly one of the mentions of *the house*. The problem is that there is no link at all in WordNet between *house* and *furniture*: the only objects mentioned as parts of houses are their structural parts, such as rooms or walls. A preliminary study of the 110 ADs in our corpus indicates that a more complex may be found in about 30-40% of the cases. This suggests limits the improvements that we can expect from more sophisticated mechanisms for searching WordNet. In fact, in our earlier work using a more complex search strategy, in less than 25% of the cases a path between a meronymic AD and the intended anchor could be found.

### 4.2 Using lexical information acquired from corpora

On the other hand, our earlier studies already revealed, first of all, that information about meronymy is perhaps the weakest aspect of WordNet's lexical base; and second, that this is one aspect of lexical knowledge where construction-based lexical acquisition methods offer the opportunity of significantly increased precision / recall ((Poesio et al., 2002a) report a 66.7% recall and 72.7% precision for this class). We tested this by a very crude strategy–given an AD with head noun N and a possible anchor with head noun N', call Google with the pattern *the N of the N'*. The results are remarkably good: at least 50 hits are found for all AD-anchor pairs in the examples in this paper, with the exception of the pair *stand-coffer* in (1f); and in all of the other examples, the correct solution is found by considering as candidate antecedents CB(S-1), CP(S-1), and FM(S-1), and choosing the one whose head noun results in the highest number of hits.

## 5 Conclusions

The first conclusion of this study is that even if we have near-perfect knowledge about focus, the first-mentioned entity of U-1 is still significanly more likely to be the anchor of a bridging description than CB(U-1). However, focusing information might still be useful provided that we found effective ways of filtering away implausible anchors, since 84.5% of anchors were previously CBs. Current versions of WordNet at least don't provide enough information for this filtering task, confirming previous results by ourselves and others (Gaizauskas et al., 1995); but corpus-based methods for acquiring information about meronymy such as those discussed in (Poesio et al., 2002a) (and other recent work) appear very promising.

## References


D. M. Carter. 1987. *Interpreting Anaphors in Natural Language Texts*. Ellis Horwood, Chichester, UK.

E. Charniak. 1986. A neat theory of marker passing. In *Proceedings of the Fifth AAAI*, pages 584–588.

H. H. Clark. 1977. Bridging. In P. N. Johnson-Laird and P.C. Wason, editors, *Thinking: Readings in Cognitive Science*. Cambridge University Press, London and New York.

R. Dale. 1992. *Generating Referring Expressions*. The MIT Press, Cambridge, MA.

C. Fellbaum, editor. 1998. *WordNet: An electronic lexical database*. The MIT Press.

K. Fraurud. 1990. Definiteness and the processing of NPs in natural discourse. *Journal of Semantics*, 7:395–433.

R. Gaizauskas, T. Wakao, K. Humphreys, H. Cunningham, and Y. Wilks. 1995. University of Sheffield: description of the LaSIE System as used for MUC-6. In *Proceedings of the Sixth Message Understanding Conference (MUC-6)*, pages 207–220. Morgan Kauffmann.

M. A. Gernsbacher and D. Hargreaves. 1988. Accessing sentence participants: The advantage of first mention. *Journal of Memory and Language*, 27:699–717.

P. C. Gordon, B. J. Grosz, and L. A. Gillion. 1993. Pronouns, names, and the centering of attention in discourse. *Cognitive Science*, 17:311–348.

B. J. Grosz, A. K. Joshi, and S. Weinstein. 1995. Centering: A framework for modeling the local coherence of discourse. *Computational Linguistics*, 21(2):202–225.

B. J. Grosz. 1977. *The Representation and Use of Focus in Dialogue Understanding*. Ph.D. thesis, Stanford.

J. A. Hawkins. 1978. *Definiteness and Indefiniteness*. Croom Helm, London.

J. Hitzeman, A. Black, P. Taylor, C. Mellish, and J. Oberlander. 1998. On the use of automatically generated discourse-level information in a concept-to-speech synthesis system. In *(ICSLP98),*, Paper 591, Australia.

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

K. R. McKeown. 1986. Discourse strategies for generating natural language text. In B. J. Grosz, K. Sparck Jones, and B. Lynn Webber, editors, *Readings in Natural Language Processing*, pages 479–500. Morgan Kaufmann, Los Altos, CA.

P. Norvig. 1989. Marker passing as a weak method for inferencing. *Cognitive Science*, 13(4):569–620.

J. Oberlander, M. O'Donnell, A. Knott, and C. Mellish. 1998. Conversation in the museum: Experiments in dynamic hypermedia with the intelligent labelling explorer. *New Review of Hypermedia and Multimedia*, 4:11–32.

M. Poesio and R. Vieira. 1998. A corpus-based investigation of definite description use. *Computational Linguistics*, 24(2):183–216, June.

M. Poesio, R. Vieira, and S. Teufel. 1997. Resolving bridging references in unrestricted text. In R. Mitkov, editor, *Proc. of the ACL Workshop on Operational Factors in Robust Anaphora Resolution*, pages 1–6, Madrid.

M. Poesio, S. Schulte im Walde, and C. Brew. 1998. Lexical clustering and definite description interpretation. In *Proc. of the AAAI Spring Symposium on Learning for Discourse*, pages 82–89, Stanford, CA, March. AAAI.

M. Poesio, F. Bruneseaux, and L. Romary. 1999. The MATE meta-scheme for coreference in dialogues in multiple languages. In M. Walker, editor, *Proc. of the ACL Workshop on Discourse Tagging*, pages 65–74.

M. Poesio, H. Cheng, R. Henschel, J. M. Hitzeman, R. Kibble, and R. Stevenson. 2000. Specifying the parameters of Centering Theory. In *Proc. of the 38th ACL*, Hong Kong, October.

M. Poesio, T. Ishikawa, S. Schulte im Walde, and R. Vieira. 2002a. Acquiring lexical knowledge for anaphora resolution. In *Proc. of the 3rd LREC*, Las Palmas, Canaria.

M. Poesio, R. Stevenson, H. Cheng, B. Di Eugenio, and J. M. Hitzeman. 2002b. A corpus-based evaluation of centering theory. NLE-TN 01, University of Essex, Department of Computer Science, April. Submitted for publication.

M. Poesio. 2000. Annotating a corpus to develop and evaluate discourse entity realization algorithms: issues and preliminary results. In *Proc. of the 2nd LREC*, pages 211–218, Athens, May.

E. F. Prince. 1992. The ZPG letter: subjects, definiteness, and information status. In S. Thompson and W. Mann, editors, *Discourse description: diverse analyses of a fund-raising text*, pages 295–325. John Benjamins.

A. J. Sanford and S. C. Garrod. 1981. *Understanding Written Language*. Wiley, Chichester.

D. Scott, R. Power, and R. Evans. 1998. Generation as a solution to its own problem. In *Proc. of the 9th International Workshop on Natural Language Generation.*

C. L. Sidner. 1979. *Towards a computational theory of definite anaphora comprehension in English discourse*. Ph.D. thesis, MIT.

M. Strube and U. Hahn. 1999. Functional centering–grounding referential coherence in information structure. *Computational Linguistics*, 25(3):309–344.

R. Vieira and M. Poesio. 2000. An empirically-based system for processing definite descriptions. *Computational Linguistics*, 26(4), December.

M. A. Walker, A. K. Joshi, and E. F. Prince, editors. 1998. *Centering Theory in Discourse*. Clarendon Press / Oxford.

Terry Winograd. 1972. *Understanding Natural Language*. Academic Press.