Squibs and Discussions

On Coreferring: Coreference in MUC and Related Annotation Schemes

Kees van Deemter*
University of Brighton

Rodger Kibble†
Goldsmiths College

In this paper, it is argued that "coreference" annotations, as performed in the MUC community for example, go well beyond annotation of the relation of coreference proper. As a result, it is not always clear what semantic relation these annotations are encoding. The paper discusses a number of problems with these annotations and concludes that rethinking of the coreference task is needed before the task is expanded. In particular, it suggests a division of labor whereby annotation of the coreference relation proper is separated from other tasks such as annotation of bound anaphora and of the relation between a subject and a predicative NP.

1. Introduction: Coreference Annotation

Various practical tasks requiring language technology including, for example, **information extraction** and **text summarization**, can be performed more reliably if it is possible to automatically find parts of the text containing information about a given topic. For example, if a text summarizer has to select the most important information, in a given text, about the 1984 Wall Street crash, then the summarization task is greatly helped if a program can automatically spot all the clauses in the text that contain information about this crash. To evaluate a program of this kind, extensive language corpora have been prepared in which human readers have annotated what has been called the **coreference** relation. These annotated corpora are then used as a gold standard against which the program's achievements can be compared. The relation of coreference has been defined as holding between two noun phrases if they "refer to the same entity" (Hirschman et al. 1997). More precisely, let us assume that α_1 and α_2 are occurrences of noun phrases (NPs) and let us assume that both have a unique referent in the context in which they occur (i.e., their context in the corpus makes them unambiguous). Under these assumptions we can use a functional notation, e.g. Referent(α), as short for "the entity referred to by α" and define (suppressing the role of context):

> **Definition**
> α_1 and α_2 *corefer* if and only if Referent(α_1) = Referent(α_2).

Putting it simply: to determine whether α_1 and α_2 corefer, first determine Referent(α_1) and Referent(α_2), then see if they are equal.

Ideally, of course, one would like to annotate many other semantic relations that hold between parts of a text, because they are also relevant for text interpretation. One candidate is the relation of **anaphora**. Loosely speaking—and glossing over some difficulties regarding the precise delimitation of anaphora (Sidner 1979; Partee 1989; van

---

* Information Technology Research Institute, University of Brighton, Lewes Road, Brighton BN2 4GJ, UK. E-mail: Kees.van.Deemter@itri.bton.ac.uk
† Mathematical and Computing Science, Goldsmiths College, University of London, London SE14 6NW, UK. E-mail: R.Kibble@gold.ac.uk

Deemter 1992)—an NP α_1 is said to take an NP α_2 as its anaphoric antecedent if and only if α_1 depends on α_2 for its interpretation (e.g., Kamp and Reyle 1993). It follows that anaphora and coreference are different things. Coreference, for example, is an equivalence relation; anaphora, by contrast, is irreflexive, nonsymmetrical, and nontransitive. Secondly, anaphora, as it has just been defined, implies **context-sensitivity** of interpretation, and this is not true for coreference. For example, a name (*President W. J. Clinton*) and a description (*Hillary Rodham's husband*) can corefer without either of the two depending on the other for its interpretation. Anaphoric and coreferential relations can coincide, of course, but not all coreferential relations are anaphoric, nor are all anaphoric relations coreferential. (An example of the latter is bound anaphora, see Section 2.1.)

Coreference annotation has been a focus of the Sixth and Seventh Message Understanding Conferences (MUC-6, MUC-7) and various other annotation exercises (e.g., Passoneau 1997; Garside, Leech, and McEnery 1997; Davies et al. 1998; Poesio 2000). In this squib, we intend to point at some fundamental problems with many of these annotation exercises, which are caused by a failure to distinguish properly between coreference, anaphora, and other, related phenomena. Because the MUC project is the best-known example of coreference annotation, on which much subsequent work is based, and because of the public availability of the MUC Task Definition (TD, Hirschman and Chinchor [1997]), we will focus on MUC.

Four criteria are listed for the MUC TD, in order of their priority (Hirschman and Chinchor 1997):

1. The MUC information extraction tasks should be supported by the annotations

2. Good (defined as ca. 95%) interannotator agreement should be achievable

3. It should be possible to annotate texts quickly and cheaply

4. A corpus should be created that can be used as a tool for linguists outside MUC.

The TD makes it clear that the annotation task has been simplified in a number of ways. For example, only NPs were annotated. Such eminently sensible simplifications notwithstanding, we will argue that the above-mentioned criteria have not been achieved and that a rethinking of the coreference annotation enterprise is in order before it ventures into new domains involving speech, noisy data, etc. (see for example, Bagga, Baldwin, and Shelton [1999]), or before it extends the relation of coreference to cover whole/part and class/instance relations (e.g. Popescu-Belis 1998; Hirschman and Chinchor 1997).

2. Problems with Coreference Annotation

In this section, some unclarities and inconsistencies will be discussed that we found in the literature on coreference annotation, and which appear to stem from confusion about what reference and coreference are. In Section 2.1, we will explore the tendency to apply coreference annotation to nonreferring NPs and bound anaphora, and we will argue that this tendency is problematic. In Section 2.2, we will argue that existing annotation enterprises still fail to respond properly to the well-known problem of how to annotate NPs that are used intensionally. In Section 2.3, we turn to a suggestion for the improvement of the actual process of annotation that has been made in the

literature, namely to separate the task of determining the "markables" from that of establishing coreference relations between them, showing that this separation is hard to maintain. At the end of each subsection, some suggestions (*Remedies*) will be made on how the problems may be tackled. These suggestions will be elaborated in Section 3.

2.1 Annotating Nonreferring NPs and Bound Anaphora

The notion of reference is common to a broad variety of semantic theories (see Gamut [1991], Chapter 1, for discussion). When speakers/writers use an NP to *refer* to an object or a set of objects, they try to single out the entity uniquely. Thus, when someone utters the NP *the tenant of the house*, the speaker may aim to single out a unique person, say Mr. *X*. Even when this is the case (i.e., the NP is used referentially rather than attributively[1]), the notion of referring has its problems. For example, the speaker may be mistaken in her belief that Mr. *X* is the tenant of the house (Donnellan 1966). In such cases it is unclear who is being referred to. Such problems notwithstanding, work on coreference annotation has usually taken the notion of reference for granted, on the assumption that clear cases, where the referent of an NP is clearly defined, outnumber the problematic ones, at least in some important types of discourse.

Let us, for now, buy into the assumption that reference is a straightforward notion. Then, following Bach (1987) (especially Sections 3.2 and 12.2), for example, one thing that is clear about reference is that some NPs do not refer. When someone says

(1)   a. *No solution emerged from our discussions*, or

        b. *Whenever a solution emerged, we embraced it*,

the *solution* NPs do not refer to any single solution, nor to any definite set of solutions. Most theorists would agree that they do not have a referent. Nonreferring NPs can enter anaphoric relations. (For example, the NP *a solution* is the (bound) anaphoric antecedent to *it* in (1b).) But if they do not refer, the *co*reference relation as defined in Section 1 (which presupposes that Referent(α_1) and Referent(α_2) are defined) is not applicable to them. Even so, the MUC TD asks annotators to treat them as if it was applicable. It acknowledges (page 10) that "one may argue that [a bound anaphor and its antecedent] are not coreferential in the usual sense," but falls short of explaining explicitly what types of anaphora are to be annotated and how (Hirschman and Chinchor 1997).[2]

The annotation of bound anaphora merits some further elaboration. Consider, for example, quantifying NPs such as *Every TV network* (or, even more problematic, *Most computational linguists* [Hirschman and Chinchor 1997], see also Section 3). If *Every TV network* refers at all, then presumably it refers to the set of all TV networks (relevant to a certain domain). The TD, however, asks annotators to let *Every TV network* corefer with *its* in (1c). According to the definition of coreference, this means that Referent(*Every TV network*) = Referent(*its*) so that Referent(*its*) is the set of all TV networks, predicting incorrectly that (1c) means (1c′):

(1)   c. *Every TV network reported its profits*

        c′. *Every TV network reported every TV network's profits*,

---

[1] See Donnellan (1966). For an interesting class of attributively used NPs, see van der Sandt (1992); examples include hypotheticals like *If this house has a tenant then the tenant is probably Dutch*, where one might ask whether *a tenant* and *the tenant* corefer.

[2] Sometimes the term "cospecification" has been employed to replace coreference (e.g. Sidner 1983; Davies et al. 1998). It is unclear, however, whether a bound anaphor and its antecedent cospecify, or how the notion should be applied to intensional constructions (Section 2.2).

(Incidentally, coreference and anaphora are not only different; they are also extremely difficult to combine into a proper equivalence relationship that allows us to recognize different clauses as being "about the same thing." Consider, for example, the relation, say R, which holds between NP_1 and NP_2 if and only if

> *either* NP_1 and NP_2 corefer (in the sense of the definition in Section 1)
> *or* NP_1 is an anaphoric antecedent of NP_2
> *or* NP_2 is an anaphoric antecedent of NP_1

Note that R is not an equivalence relation. The subject of (1c), for example, can corefer with a plural pronoun in the next sentence, e.g., (...) They *are now required to do this*, but *They* and *it* (in (1c)) do not stand in the relation R.)

Predicative NPs are another category of NPs whose referentiality is problematic and yet the MUC TD instructs annotators to let them corefer with other NPs. In (2a) and (2b), for example, the predicative NP *the/a president of DD* cannot be replaced by the proper name *Higgins* without changing the meaning of the sentence beyond recognition, indicating that the relation between the two NPs must be something other than coreference:

(2)   a. *Higgins was/became the/a president of DD*

        b. *Higgins, once the president of DD, is now a humble university lecturer*

We will have more to say about predicative NPs in the next section.

To sum up, MUC's annotators have been instructed to let NPs of all major classes (definite, quantificational, and indefinite) "corefer" liberally with other NPs, even when it is far from clear that the NPs in question have been used referentially. As a result, the relation actually annotated in MUC—henceforth called the IDENT relation, following Hirschman and Chinchor (1997)—must be distinguished from the coreference relation. The TD admits that certain instructions may be incompatible with the definition of coreference but no reason is given for these incompatibilities and no intuitive motivation for the relation IDENT is offered. As a result, the annotator is left with a long series of instructions that fail to be held together by a common rationale.

Remedy. Go back to basics: start from a definition of coreference and write a TD that implements the definition. We suggest that it is not until this has been done successfully that extensions into the area of bound anaphora become a risk worth taking.

2.2 Problems of Intensionality and Predication

Problems posed to coreference annotation by intensionality (Hirschman et al. 1997) have motivated considerable complications in the TD. Consider Section 6.4, which discusses the implications of "change over time." The TD says that "two markables should be recorded as coreferential if the text asserts them to be coreferential at ANY TIME" (Hirschman and Chinchor 1997, page 11). Thus, for example, the TD points out that in a case like

(3)   *Henry Higgins, who was formerly sales director of Sudsy Soaps, became president of Dreamy Detergents*

annotators are expected to mark (1) *Henry Higgins*, (2) *sales director of Sudsy Soaps*, and (3) *president of Dreamy Detergents* as coreferential. (Similar strategies seem to be

adopted by most practitioners of coreference annotation, e.g., Cristea et al. [1999]). But since coreference is generally agreed to be a equivalence relation (e.g. Hirschman and Chinchor 1997, Section 1.3), this implies that the sales director of Sudsy Soaps and the president of Dreamy Detergents are the same person. Clearly, this cannot be right.

Luckily, there are other parts of the same TD that do a better job of applying the notion of coreference to sentences involving change over time. Consider, for example, Section 1.3, where the sentence *the stock price fell from $4.02 to $3.85* is discussed. Here annotators are asked to consider *the stock price* as standing in the IDENT relation with *$3.85* but *not* with *$4.02*, because *$3.85* is "the more recent value" (p. 3). (If both coreferred with *the stock price*, it would have followed that $4.02 and $3.85 are equal.) This solution, however, is still problematic. What, for instance, if the price continues to fall?

(4)   a. *The stock price fell from $4.02 to $3.85*;

        b. *Later that day, it fell to an even lower value, at $3.82.*

Does the annotator have to go back to (4a), deciding that $3.82 is an even more recent value and *the stock price* does not stand in the IDENT relation with $3.85 after all?

Remedy. At least three different strategies are conceivable. Perhaps most obviously, one might decide that coreference between a functional description like those in (3) or (4) and an NP denoting a value requires this value to be the present (rather than the most recent) value of the function. But, the text does not always say what the present value is. Moreover, functional descriptions do not always pertain to the present. In *Last year, the president resigned*, for example, the subject refers to *last year*'s president, and consequently, it does not corefer with NPs referring to the present president. A second strategy, consistent with Dowty, Wall, and Peters (1981, Appendix iii) might be to say that *The stock price* refers only to a Montague-type individual concept, that is, a function from times to numbers. It would follow that *The stock price* does not corefer with either *$4.02* or *$3.85* and no problem would arise. Analogously, *president of Dreamy Detergents*, in (3) above, where it is used predicatively, might denote an individual concept rather than an individual. If the next sentence goes on to say *He died within a week*, then *he* is coreferential with *Henry Higgins*; if, instead, the text proceeds *This is an influential position, but the pay is lousy*, then *This* is coreferential with *president of Dreamy Detergents*. If both these analyses prove to be too complex to be used in large-scale annotation exercises, one might have to take the point of view that such descriptions simply do not refer. This would amount to a third strategy, which excludes these descriptions from entering coreference relations altogether and leaving their analysis to the other tasks.

2.3 What's Markable?
It has been proposed that annotation can profitably be broken down into two more manageable steps: annotation of markables (step 1) is to be carried out before (step 2) partitioning the set of markables into equivalence classes of coreferring elements (e.g., Hirschman and Chinchor 1997). It turns out, however, that a strict distinction between the two steps is difficult to maintain, because, in principle, almost anything is markable. In the MUC-7 TD, this is sensibly acknowledged by letting annotators mark up certain elements only if they corefer with an existing markable: these include conjuncts and prenominal modifiers. In the following example, the first occurrence of *aluminum* is only considered to be markable because it corefers with the occurrence of this noun

as a bare NP in the second clause.

(5)   *The price of* aluminum *siding has steadily increased, as the market for*
        aluminum *reacts to the strike in Chile.* (Hirschman and Chinchor 1997)

In other words: coreference (step 2) helps to determine what the markables are (step 1).
Finding all the NPs that *might* participate in coreference becomes even harder if the
annotation scheme is extended to cover event coreference (noted in the "wish list" in
Section 1.4 of the TD) since it is often extremely difficult to determine which events
can serve as antecedents (Hirschman and Chinchor 1997):

(6)   *Be careful not to get the gel in your eyes. If* this *happens, rinse your eyes with*
        clean water and tell your doctor. (ABPI, 1997)

Examples of this kind suggest that one markable (e.g., an event) can give rise to
another (e.g., the negation of the event). A complication of a similarly algebraic flavor
arises if "discontinuous elements, including conjoined elements" are covered, as when
a plural pronoun corefers with a combination of previously occurring NPs (Hirschman
and Chinchor 1997, section 1.4; see Garside, Leech, and McEnery [1997] for a proposal).
Note especially that annotators would have to be on guard for the possibility of *different*
combinations of markables coreferring to each other. A corpus, for example, can easily
contain NPs A, B, C, and D for which Referent(A) $\cup$ Referent(B) $=$ Referent(C) $\cup$
Referent(D). Even assuming that each of A, B, C, and D has been properly identified
as a markable during step 1, this is little guarantee that annotators of step 2 will realize
the complex coreference relation between the combination of A and B and that of C
and D. (Recall that coreference relations are to be annotated even in the absence of an
anaphoric relationship.) The number of possible combinations of markables (some 2^n
when there are n markables) will often be too large to handle.

Remedy. One alternative is to have a first pass where only referring expressions that
look like anaphors are marked up, such as pronouns and definite NPs. Subsequent
passes would look for antecedents for these expressions and link coreferring elements.
An intermediate approach would be to mark up a core set of referring expressions on
the first pass, allowing for further referring expressions to be identified on subsequent
passes if this is necessary to resolve coreference. The extent to which each strategy
would contribute to accuracy and speed of annotation remains to be determined.

3. Conclusion

Current "coreference" annotation practice, as exemplified by MUC, has overextended
itself, mixing elements of genuine coreference with elements of anaphora and predica-
tion in unclear and sometimes contradictory ways. As a result, the annotated corpus
emerging from MUC is unlikely to be as useful for the computational linguistics re-
search community as one might hope (Criterion 4, see Section 1), the more so because
generalization to other domains is bound to make problems worse. In many domains,
for example, other sources of intensionality than change over time occur prominently.
An example is epistemic modality:

(7)   *Henry Higgins might be the man you have talked to.*

The relation between *Henry Higgins* and *the man you have talked to* is analogous to that between *Henry Higgins* and *sales director of Sudsy Soaps* in (3), with possible worlds taking the place of points in time: the two NPs refer to the same individual in some possible worlds only (see Groenendijk, Stokhof, and Veltman [1996] for relevant theoretical work). Modality, of course, interacts with tense (as in *Henry Higgins might become the president of Dreamy Detergents*), leading to further complications.

The MUC TD has addressed many of the difficult problems in the area of reference and coreference, but if its success is judged by the criteria in Hirschman and Chinchor (1997) (see Introduction), the results are mixed at best. Criterion 4 has been discussed above. Concerning Criterion 3, it appears doubtful that the present task definition can be applied "quickly and cheaply." Hirschman et al. (1997), when discussing this issue, note that interannotator agreement, at the time of writing, was in the low eighties. This figure, which falls markedly short of the 95% required by Criterion 2, does not seem to have improved substantially since (Breck Baldwin, personal communication). Concerning Criterion 1, finally, it has been observed that the figures for recall in the MUC information extraction algorithm are rather discouraging (Appelt 1998). The material in Section 2 suggests that this relative lack of success is no accident and that unclarities in the TD are to blame. Repairs are not always easy to find. Given this situation, we suggest that a rethinking of the coreference task is required.

Firstly, one needs a consistent story of what reference and coreference are taken to be. Theoretical work on reference does not show a consensus on some crucial questions in this area (Bach 1987; Kronfeld and Roberts 1998). Different answers have been suggested, each with its own advantages and disadvantages. For example, one might identify the notion of a *referring* NP with that of a *semantically definite* NP in the sense of Barwise and Cooper (1981).[3] This would include proper names, extensional definite descriptions, universally quantified NPs, and specifically used indefinites (e.g., *a company whose name is withheld*), but it would exclude nonspecifically used indefinites such as *at least n companies, most computational linguists*. A more liberal approach along the lines of Kamp and Reyle (1993, Chapter 4), would predict that a quantifying NP such as the subject of *Most computational linguists use a parser* refers to the set of those computational linguists *who use a parser*: the VP helps to determine the referent of the NP. The first approach would make annotation easier to perform and the results would be likely to be more reliable as a result, but it would feed less information into the information extraction task. Trade-offs of this kind are unavoidable, and experimentation will be required to determine which option provides the best results.

Secondly, we suggest a further division of labor whereby those phenomena that are no longer accounted for in the new TD are covered by other tasks (Kibble and van Deemter 2000). For example, the two NPs *Henry Higgins* and *president of Sudsy Soaps* (example (3)) do not corefer, and the relation between them should be irrelevant to coreference annotation. If it is imperative that information about Henry's previous jobs be saved for posterity then some other annotation task has to be defined, with its own very different TD, involving the notion of individuals having properties at certain times or intervals only. Something analogous is true for the annotation of bound anaphora.

The issue under discussion illustrates a more general point. It is now widely agreed that linguistic theorizing is sometimes insufficiently informed by observational data. Conversely, we would like to submit that corpus-based research is sometimes

---

3 A semantically definite NP α is one whose set-theoretic denotation takes the form of a principal filter (Partee, ter Meulen, and Wall 1990), i.e., a set of the form $\{X: Y \subseteq X\}$ for some set of individuals Y.

insufficiently informed by theory. It follows, in our opinion, that there is scope for more collaboration between theoretical and corpus-based linguists in this area. This squib attempts to be a small step in this direction.

References

ABPI. 1997. *1996–1997 ABPI Compendium of Patient Information Leaflets*. Association of the British Pharmaceutical Industry.

Appelt, Douglas. 1998. An overview of information extraction technology and its application to information retrieval. In *Proceedings of TWLT14, Language Technology in Multimedia Information Retrieval*, pages 49–58, Twente.

Bach, Kent. 1987. *Thought and Reference*. Clarendon Press, Oxford.

Bagga, Amit, Breck Baldwin, and Sara Shelton. 1998. Coreference and its applications. Call for papers for workshop associated with the 37th Annual Meeting of the Association for Computational Linguistics, University of Maryland, 1999. See www.cs.duke.edu/~amit/acc99-wkshp.html.

Barwise, Jon and Robin Cooper. 1981. Generalized quantifiers and natural language. *Linguistics and Philosophy*, 4:159–219.

Cristea, Dan, Nancy Ide, Daniel Marcu, and Valentin Tablan. 1999. Discourse structure and coreference: An empirical study. In Dan Cristea, Nancy Ide, and Daniel Marcu, editors, *Proceedings of ACL'99 Ws: The Relation of Discourse/Dialogue Structure and Reference*, pages 46–53.

Davies, Sarah, Massimo Poesio, Florence Bruneseaux, and Laurent Romary. 1998. Annotating coreference in dialogues: Proposal for a scheme for MATE. See www.cogsci.ed.ac.uk/~poesio/MATE/anno-manual.html.

Donnellan, Keith. 1966. Reference and definite descriptions. *Philosophical Review*, 75:281–304.

Dowty, David, Robert Wall, and Stanley Peters. 1981. *Introduction to Montague Semantics*. Kluwer, Dordrecht.

Gamut, L. T. F. 1991. *Logic, Language and Meaning*, Volume 2. University of Chicago Press, Chicago.

Garside, Roger, Geoffrey Leech, and Tony McEnery. 1997. *Corpus Annotation*. Longman, London.

Groenendijk, Jeroen, Martin Stokhof, and Frank Veltman. 1996. Coreference and modality. In Shalom Lappin, editor, *The Handbook of Contemporary Semantic Theory*. Blackwell, Cambridge, MA, pages 179–214.

Hirschman, Lynette and Nancy Chinchor. 1997. MUC-7 coreference task definition. In *MUC-7 Proceedings*. Science Applications International Corporation. See www.muc.saic.com.

Hirschman, Lynette, Patricia Robinson, John Burger, and Marc Vilain. 1997. Automating coreference: The role of annotated training data. In *Proceedings of AAAI Spring Symposium on Applying Machine Learning to Discourse Processing*.

Kamp, Hans and Uwe Reyle. 1993. *From Discourse to Logic*. Kluwer, Dordrecht.

Kibble, Rodger and Kees van Deemter. 2000. Coreference annotation: Whither? In Maria Gavrilidou et al., editors, *Proceedings of the 2nd International Conference on Language Resources and Evaluation*, pages 1,281–1,286, Athens.

Kronfeld, Ami and Lawrence Roberts. 1998. *Special Issue on Reference, Pragmatics and Cognition 6*. John Benjamins, Amsterdam and Philadelphia.

Partee, Barbara. 1989. Binding implicit variables in quantified contexts. *Proceedings of the Chicago Linguistic Society*, 25:342–365.

Partee, Barbara, Alice ter Meulen, and Robert Wall. 1990. *Mathematical Methods in Linguistics*. Kluwer, Dordrecht.

Passoneau, Rebecca. 1997. Instructions for applying discourse reference annotation for multiple applications (DRAMA). Unpublished manuscript.

Poesio, Massimo. 2000. Annotating a corpus to develop and evaluate discourse entity realization algorithms: Issues and preliminary results. In Maria Gavrilidou et al., editors, *Proceedings of the 2nd International Conference on Language*

Resources and Evaluations, pages 211–218, Athens.

Popescu-Belis, Andrei. 1998. How corpora with annotated coreference links improve reference resolution. In Antonio Rubio et al., editors, *First International Conference on Language Resources and Evaluation*, pages 567–572, Granada. European Language Resources Association.

Sidner, Candace. 1979. *Towards a Computational Theory of Definite Anaphora Comprehension in English Discourse*. Ph.D. dissertation, AI Lab, MIT, Cambridge, MA.

Sidner, Candace. 1983. Focusing in the comprehension of definite anaphora. In Michael Brady and Robert Berwick,

editors, *Computational Models of Discourse*. MIT Press, Cambridge, MA, pages 267–330.

van Deemter, Kees. 1992. Towards a generalization of anaphora. *Journal of Semantics*, 9:27–51.

van Deemter, Kees and Rodger Kibble. 1999. What is coreference and what should coreference annotation be? In Amit Bagga, Breck Baldwin, and Sara Shelton, editors, *Proceedings of ACL workshop on Coreference and Its Applications*, pages 90–96, Maryland.

van der Sandt, Rob. 1992. Presupposition projection as anaphora resolution. *Journal of Semantics*, 9:333–37.