Event versus entity co-reference:
Effects of context and form of referring expression

Sharid Loáiciga[1]   **Luca Bevacqua**[2]   **Hannah Rohde**[2]   **Christian Hardmeier**[3]

[1]CLASP, University of Gothenburg
[2]Department of Linguistics and English Language, University of Edinburgh
[3]Department of Linguistics and Philology, Uppsala University

`sharid.loaiciga@gu.se`   `lbevacqu@ed.ac.uk`
`hannah.rohde@ed.ac.uk`   `christian.hardmeier@lingfil.uu.se`

Abstract

Anaphora resolution systems require both an enumeration of possible candidate antecedents and an identification process of the antecedent. This paper focuses on (i) the impact of the form of referring expression on entity-vs-event preferences and (ii) how properties of the passage interact with referential form. Two crowd-sourced story-continuation experiments were conducted, using constructed and naturally-occurring passages, to see how participants interpret *It* and *This* pronouns following a context sentence that makes available event and entity referents. Our participants show a strong, but not categorical, bias to use *This* to refer to events and *It* to refer to entities. However, these preferences vary with passage characteristics such as verb class (a proxy in our constructed examples for the number of explicit and implicit entities) and more subtle author intentions regarding subsequent re-mention (the original event-vs-entity re-mention of our corpus items).

1 Introduction

A challenge in discourse interpretation is the resolution of referring expressions, particularly those whose meaning is compatible with many potential antecedents. To take an example like (1), a passage may introduce a number of entities and situations that a subsequent sentence might refer to.

(1) <u>Everybody</u> who is involved with <u>this debate</u> has been struggling over <u>me</u> and <u>my personality</u>. [ParCorFull]

For a sentence following (1), certain expressions would be resolved unambiguously to a unique entity (e.g., to the speaker for a 1st person singular pronoun *I*) or would easily be linked to the only compatible referent in the context (e.g., to the group of relevant individuals described as *Everybody* for a 3rd person plural pronoun *They*). Other expressions are compatible with more than one entity (e.g., the debate or the personality for a pronoun *It*) and therefore create potential ambiguity. Making matters worse, the antecedent of some expressions can be either an entity or something more abstract: an event or situation or idea. Such expressions include personal pronouns like *It* and demonstrative pronouns like *This/That*.

Given the complexity of identifying a set of candidate abstract antecedents in a given context and then determining whether a particular expression is re-mentioning one of those abstract antecedents or a more concrete entity, many co-reference systems focus only on nominal antecedents (e.g., BART, Stanford's sieve-based, HOTCoref (Versley and Björkelund, 2015)). However, event instances are also referential.[1]

This paper asks when and to what degree event instances serve as antecedents when a competing entity referent is also available. The goal is to model human choices as a baseline to inform co-reference systems. We report two psycholinguistic studies that use a story-continuation task to measure participants' resolution of pronouns *It* and *This*.

Improving our understanding of the interpretation of the "difficult" anaphoric cases is a step towards better anaphora and co-reference systems. It has been noted that current systems struggle to identify this type of reference and that anaphoricity determiners have poor performance (Heinzerling et al., 2017). *It*, *This* and *That* are also frequent in dialogue data for which co-reference sys-

---

[1] Here we call *event* what is more commonly known as *abstract anaphora* (cf. Dipper and Zinsmeister (2010); Nedoluzhko and Lapshinova-Koltunski (2016)). We take as an *event* any non-nominal relationship for the pronouns *It* and *This* and a *textual* antecedent in the form of a text span of variable length (e.g., a word, a clause, several sentences). *Textual* means that anaphoric relations for which some type of inference is necessary are not included, e.g., bridging or extra-textual reference. The term *event reference* is founded upon Webber (1986), and we set on the name *event* for the sake of consistency with the annotation in the corpus used in the second study presented here.

tems' performance is particularly low (Eckert and Strube, 2000; Müller, 2007). In addition, pronoun function is relevant to the evaluation of machine translation systems since different functions entail different translations according to the constraints of the language pair and can thus affect performance (Guillou, 2016).

2 Related Work

Both corpus-based and psycholinguistics works on the interpretation of anaphoric expressions concentrate on the identification of the antecedents of nominal expressions. Abstract anaphora—anaphora that involve reference to abstract entities such as events or states (Asher, 1993)—is much less studied from both fields, as evidenced by the little amount of annotated data available (Dipper and Zinsmeister, 2010; Poesio, 2015).

Corpus-based studies of pronouns are often done in relationship to the texts on which co-reference resolution systems will be trained and tested. With the clear aim to improve precision, the authors of these systems have an interest in quantifying "non-anaphoric" pronouns for preventing their resolution. We know for instance, that about 5% of referential pronouns and 71% of demonstratives in dialogue data refer to events (Müller, 2007; Poesio, 2015), whereas about 3% of referential *it* pronouns in written text of various genres refer to events (Evans, 2001).

In psycholinguistic research, on the other hand, the focus has been on using theoretical constructs of complexity, salience, and focus to capture co-reference patterns. The demonstratives *This* and *That* have been grouped together, assuming that they behave in the same manner, but potentially differently from *It*. Brown-Schmidt et al. (2005) analyze *It* vs *That* and report a preference for *That* if what is referred to is a composite (e.g., *I'll have the hamburger and fries. I'll have that, too.*), independent of other metrics of the salience of the referent. Building on the Centering co-reference model (Grosz et al., 1995), Passonneau (1989) analyzes intra-sentential instances of *It* vs *That* with an explicit NP antecedent. She reports that *It* is used to refer to the center (most often the subject), whereas *That* favors non-centers.

Corpus-based studies offer insights about language use, since the written texts they are based on are, after all, natural passages. They offer better estimates for building systems that will be used

on those texts. Corpus-based studies, on the other hand, do not offer any explanation as to why a particular item follows a certain distribution, and they grant little control over the confounding variables responsible for that distribution. In this respect, psycholinguistics studies provide more suitable methods for capturing the cognitive processes behind naturally occurring phenomena. We therefore start the next section with a study using constructed passages to allow for careful control over format and content.

3 Study 1: Constructed passages

A story-continuation experiment was conducted to establish a baseline rate at which participants assign *It/This* pronouns to entity vs event antecedents. By varying a property of the context sentence, we test how malleable the two pronouns' respective co-reference preferences are. A 2x2 design manipulated the context sentence (alternating/non-alternating verb) and the pronoun prompt (*It/This*, as in (2)-(3)).

(2) The train from the Highlands arrived promptly. It/This ____

(3) The balloon with the red hearts popped noiselessly. It/This ____

The availability of entities for anaphoric resolution is dependent on the argument structure of the previous predicates. Alternating verbs can have an intransitive as well as a transitive use: the first usually describes a change of state (4-a), and the latter specifies, in subject position, which entity brought on the change (4-b). Conversely, non-alternating verbs do not allow a transitive use (5).[2]

(4) a. The snow melted.
    b. The heat melted the snow.

(5) a. The battery died.
    b. * The heat died the battery.

Manipulating the verb in the context sentence affects the argument realization options associated with the predicate: Non-alternating verbs like *arrive* permit only a single realization with the entity that arrives always in subject position; alternating verbs like *pop* are compatible with realizations

---

[2] Jespersen (1927) collects verbs undergoing alternation in a "move and change class". They have also been referred to as respectively *causative* and *anticausative* (or *inchoative*) verbs (Schäfer, 2009); the phenomenon has also been studied as "causative-inchoative alternation" (Haspelmath, 1993).

where the entity that pops appears in subject position or object position. For alternating verbs, an explicit agent entity can be introduced (*I popped the balloon*) or left implicit, as in (3).

One hypothesis is that alternating verbs could make available an additional (implicit) agent who might provide more entity co-reference opportunities and thereby increase entity co-reference and reduce event co-reference. Another hypothesis is that non-alternating verbs may make salient one single (explicit) entity by eliminating competition from other (implicit) entities and thereby yield more entity co-reference and less event co-reference. The existence of an external, unspecified argument in the syntax of alternating verbs is still controversial (Embick, 2004; Schäfer, 2009), but the cognitive accessibility of a possible agentive entity arises from the very fact that the causative alternation exists.

Although differences have been observed between the use of proximal and distal demonstratives *this* and *that* (Çokal et al., 2014), we targeted only one demonstrative pronoun in order to simplify the design. This is in keeping with observations about the functional grouping of a number of pronouns (zeros, demonstratives, and personal pronouns) when used deictically (Webber, 1990).

3.1 Materials

The 24 experimental items consisted of a context sentence and a pronoun prompt, as in (2)-(3). Participants saw all items, with either *It* or *This*. Subject NPs were modified (8 nouns with pre-nominal adjectives, 8 nouns with post-nominal prepositional phrases, 8 nouns with post-nominal relative clauses). The verb used an adverbial or particle predicate (roughly half alternating, half non-alternating). The head of the subject NP was always the only singular entity, with any other mentioned entities being incompatible with 3rd person singular co-reference (e.g., *we* or *the red hearts*).[3]

The 24 experimental items were interleaved with 40 filler items. These included 20 passages with a context sentence mentioning one or two entities, followed by a discourse adverbial prompt (e.g., *As a result*, *Then*), 16 passages for an unrelated experiment involving mentions of companies and other organisations, and 4 catch trials with an obvious correct response (e.g., *Caleb did all the*

---

[3] The data to reproduce our experiments and the full models can be found on `https://github.com/sharidloaiciga/event_vs_entity`.

cooking for the BBQ even though he hates BBQ. He prefers mac 'n ___).

3.2 Participants

Twenty-seven monolingual English-speaking participants aged 19-63 (mean age 36, σ=11.2; 15 male) were recruited from Amazon's Mechanical Turk (Munro et al., 2010; Gibson et al., 2011) and received $4 for an estimated 30-minute task.

3.3 Procedure

Continuations were collected via a web-based interface that participants could access from their own computer. Each item was presented on a page by itself with a text box for participants to use for writing their continuation.

3.4 Annotation and analysis

Continuations for experimental items were annotated for type of co-reference (entity vs event). The four authors of this paper shared the annotation such that all target continuations were coded by two annotators. To be conservative, annotators were blind to the *It/This* prompt condition and agreed to err on the side of annotating a pronoun as ambiguous if the pronoun could be interpreted plausibly as coreferential with an event or an entity (e.g., *The brand new siren sounded loud.* [*omitted pronoun*] *startled some people*).

Using mixed-effects logistic regression, we modeled the binary outcome of entity or event co-reference with fixed effects for prompt type, verb class, and their interaction, with maximal random effects structure when supported by the data (Barr et al., 2013). Where a model did not converge, we removed random correlations. All factors were centered. Reported p-values are from glmer model output using the lme4 package (Bates et al., 2015) in R (R Development Core Team, 2008).

3.5 Results

Of the 626 total continuations, we excluded 128 that were judged by one or more annotators to be ambiguous (or for which the annotators gave conflicting annotations) as well as 55 that used the prompt in another way (e.g., *This noun*). This left 443 continuations with either entity or event co-reference. Note that at the analysis stage, 2 of the 24 verbs were re-classified as alternating verbs, shifting the original even split between alternating/non-alternating verbs. However,

glmer models are understood to be robust against datasets that are not perfectly balanced.

The results (see Figure 1) show a strong, but not categorical, bias to use *It* to refer to entities and *This* to refer to events. In addition, verb type impacts co-reference, whereby verbs that permit alternations yield more event co-reference than non-alternating verbs. This is in keeping with our second hypothesis that the salience of the single argument of non-alternating verbs may have attracted more entity co-reference.
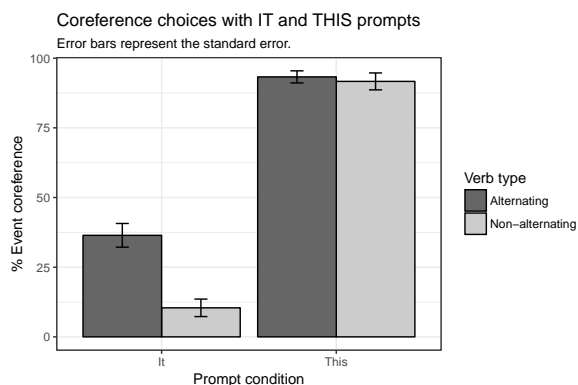

Coreference choices with IT and THIS prompts
Error bars represent the standard error.

Figure 1: Study 1 results by prompt and verb type.

The prompt type × verb type model of co-reference choice confirms a main effect of prompt type (β=5.100, p<0.001) and a main effect of verb type (β=1.437, p<0.05). There was no prompt × verb type interaction (β=-1.350, p=0.22).[4]

4 Study 2: Corpus passages

4.1 Materials

The 48 target passages are minimally edited sentences extracted from the ParCorFull corpus (Lapshinova-Koltunski et al., 2018). This is a German-English parallel corpus annotated with full co-reference. Although the corpus is designed for nominal co-reference, it includes annotations of two types of antecedents: entities and events. Entities can be either pronouns or NPs, whereas events can be VPs, clauses or a set of clauses.

ParCorFull includes texts from TED talks transcripts and also newswire data.[5] Since pronouns

are generally more frequent in the TED talks genre than news, we concentrated on this portion of the corpus only. Twelve examples of each *It-entity*, *It-event*, *This-entity*, and *This-event* were selected. In comparison to the sentences from Study 1, the corpus sentences were relatively long; therefore, simplified or shortened versions were used.

Additionally, the target passages were interleaved with 52 filler items. From these, 24 were extracted from ParCorFull sentences with no annotation and a continuation starting with an adverbial expression was prompted (e.g., *The encyclopedia business in the days of leatherbound books was basically a distribution business. Eventually, ____*). 24 other fillers were extracted from the OntoNotes corpus (Pradhan et al., 2013) for a dataset for an unrelated experiment involving mentions of companies and other organisations, as in Study 1. A final 4 fillers repeated the catch trials from Study 1.

4.2 Participants

Nineteen monolingual English-speaking participants aged 23-44 (mean age 30, σ=6.5; 13 male) were recruited from Amazon's Mechanical Turk and received $7 for an estimated 50-minute task.

4.3 Procedure, annotation, and analysis

The procedure was identical to that in Study 1. The annotation followed that described for Study 1. As an illustration, example (6) shows a passage whose original co-reference relation was one between an *it* pronoun and an entity antecedent. The continuations in (7) were annotated as event co-reference (7-a), entity co-reference (7-b), and no co-reference when the *It* prompt was classed as being used pleonastically (7-c).

(6) You carry a phone. It knows where you are. [original co-reference: entity∼*it*]

(7) a. You carry a phone. This is something that just about everyone does these days.
   b. You carry a phone. It is capable of connecting you to others and the world around you.
   c. You carry a phone. It wouldn't hurt you to call once in a while.

The binary outcome of entity/event co-reference

---

[4] Inspection of Figure 1 suggests a possible interaction whereby the effect of verb type looks stronger in the *It* condition than in the *This* condition. The lack of a significant interaction in the model may reflect the fact that the co-reference rate for non-alternating verbs in the *This* condition is already near ceiling and there may be little room for (measuring) a further increase.

[5] Specifically, the ParCorFull corpus includes the datasets

used in the ParCor corpus (Guillou et al., 2014), the DiscoMT workshop (Hardmeier et al., 2016) and the test sets from the WMT 2017 shared task (Bojar et al., 2017).

was again modeled with a logistic regression. We included fixed effects for prompt type, original passage co-reference (entity/event), original passage referring expression (*it/this*), and the 2-way and 3-way interactions. All factors were centered.

4.4 Results

Of the 788 total continuations, we excluded 94 that were judged by one or more annotators to be ambiguous (or for which the annotators gave conflicting annotations) as well as 98 that used the prompt in another way (e.g., *This noun*). This left 596 continuations with either entity or event co-reference.
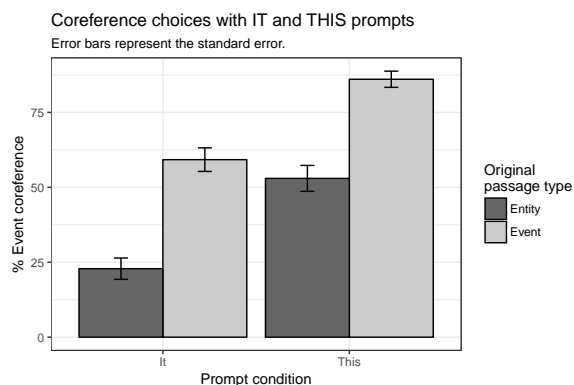

Figure 2: Study 2 results by prompt and original co-reference (collapsing over original *it/this* pronoun type)

The results (see Figure 2) follow those of Study 1 for the prompt manipulation: Event co-reference is higher with *This* than *It*. Event co-reference further increases when the original passage contained event co-reference. The model (prompt type × original passage type × original passage pronoun) confirms a main effect of prompt type (β=2.529, p<0.001) and a main effect of original passage type (β=3.053, p<0.001), with no effect of original pronoun or any significant interactions.

5 Discussion

The two studies show divergent co-reference distributions for the personal pronoun *It* and the demonstrative *This*: a bias towards entity co-reference for *It* and a bias for event co-reference for *This*. As far as we know, this pattern has been proposed (Dipper and Zinsmeister, 2010), but not properly measured. Given the oft-assumed division of labor between these two pronouns, what is notable is their flexibility. Neither form was found to be used categorically in Study 1 or Study 2.

Interestingly, the study with the constructed passages showed that verbs which permit an agent alternation as either an implicit or explicit argument are more prone to trigger an event co-referent than an entity one. This finding is potentially useful as an additional feature for anaphoricity detection or event mention identification in co-reference resolution systems.

Furthermore, we saw a bias towards event co-reference for the corpus passages in Study 2 that were known to have yielded event co-reference in their original passages. This suggests that there are properties of the context sentence that may make salient an event over an entity. If there are event-favoring properties of the context sentence that human participants are sensitive to, it is a tractable task to build automatic classifiers that learn to recognize such properties. This supports the idea that the task of differentiating anaphoric and pleonastic instances of *It* (Evans, 2001; Boyd et al., 2005; Bergsma and Yarowsky, 2011; Lee et al., 2016; Loáiciga et al., 2017) could potentially improve performance.

Although presumably (machine) learnable, the question of what exactly constitutes an event remains unanswered. A number of ambiguous examples which were excluded from our analysis included entities that are close to their entailed event (e.g., *The bomb that the arsonists had planted exploded violently*) or that were very abstract (e.g., *The greatest opportunity materialized unexpectedly. It/This was almost like magic.*).

6 Conclusions and Future Work

This paper reports an investigation on abstract anaphora. Specifically, two studies targeted the ambiguity that occurs when entity and event antecedents are available for the pronouns *It* and *This*. A clear pattern emerged whereby *It* favors entity co-reference and *This* favors event co-reference. This pattern is also affected by the number of arguments that the main verb can take. Although further investigation is needed regarding the properties of events, their salience, and the gray area between events and entities, our results take a first step towards disentangling the behavior of less well-understood anaphoric relations.

References

Nicholas Asher. 1993. *Reference to Abstract Objects in Discourse*. Springer, Netherlands.

Dale J. Barr, Roger Levy, Christoph Scheepers, and Harry J. Tily. 2013. Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of memory and language*, 68(3):255–278.

Douglas Bates, Martin Mächler, Ben Bolker, and Steve Walker. 2015. Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, 67(1):1–48.

Shane Bergsma and David Yarowsky. 2011. NADA: A robust system for non - referential pronoun detection. In Iris Hendrickx, Sobha Lalitha Devi, António Branco, and Ruslan Mitkov, editors, *Anaphora Processing and Applications: 8th Discourse Anaphora and Anaphor Resolution Colloquium (DAARC)*, Lecture Notes in Artificial Intelligence, pages 12–23. Springer, Faro, Portugal.

Ondřej Bojar, Rajen Chatterjee, Christian Federmann, Yvette Graham, Barry Haddow, Shujian Huang, Matthias Huck, Philipp Koehn, Qun Liu, Varvara Logacheva, Christof Monz, Matteo Negri, Matt Post, Raphael Rubino, Lucia Specia, and Marco Turchi. 2017. Findings of the 2017 conference on machine translation (wmt17). In *Proceedings of the Second Conference on Machine Translation, Volume 2: Shared Task Papers*, pages 169–214, Copenhagen, Denmark. Association for Computational Linguistics.

Adriane Boyd, Whitney Gegg-Harrison, and Donna K. Byron. 2005. Identifying non-referential *it*: a machine learning approach incorporating linguistically motivated patterns. In *Proceedings of the ACL Workshop on Feature Engineering for Machine Learning in Natural Language Processing*, pages 40–47, Ann Arbor, Michigan. Association for Computational Linguistics.

Sarah Brown-Schmidt, Donna K. Byron, and Michael K. Tanenhaus. 2005. Beyond salience: Interpretation of personal and demonstrative pronouns. *Journal of Memory and Language*, 53(2):292–313.

Derya Çokal, Patrick Sturt, and Fernanda Ferreira. 2014. Deixis: *This* and *That* in written narrative discourse. *Discourse Processes*, 51(3):201–229.

Stefanie Dipper and Heike Zinsmeister. 2010. Towards a standard for annotating abstract anaphora. In *Proceedings of the LREC Workshop on Language Resource and Language Technology Standards state of the art, emerging needs, and future developments*, LREC10-W4, pages 54–59, Valleta, Malta. European Language Resources Association (ELRA).

Miriam Eckert and Michael Strube. 2000. Dialogue acts, synchronising units and anaphora resolution. *Journal of Semantics*, 17(1):51–89.

David Embick. 2004. Unaccusative syntax and verbal alternations. In Artemis Alexiadou, Elena Anagnostopoulou, and Martin Everaert, editors, *The Unaccusativity Puzzle: Explorations of the Syntax-Lexicon Interface*, pages 137–158. Oxford University Press, Oxford, New York.

Richard Evans. 2001. Applying machine learning toward an automatic classification of it. *Literary and Linguistic Computing*, 16(1):45–57.

Edward Gibson, Steve Piantadosi, and Kristina Fedorenko. 2011. Using mechanichal turk to obtain and analyze english acceptability judgments. *Language and Linguistics Compass*, 5(8):509–524.

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

Liane Guillou. 2016. *Incorporating Pronoun Function into Statistical Machine Translation*. Ph.D. thesis, University of Edinburgh, Scotland, UK.

Liane Guillou, Christian Hardmeier, Aaron Smith, Jörg Tiedemann, and Bonnie Webber. 2014. ParCor 1.0: A parallel pronoun-coreference corpus to support statistical MT. In *Proceedings of the 9th International Conference on Language Resources and Evaluation*, LREC 2014, pages 3191–3198, Reykjavik, Iceland. European Language Resources Association (ELRA).

Christian Hardmeier, Jörg Tiedemann, Preslav Nakov, Sara Stymne, and Yannick Versely. 2016. DiscoMT 2015 Shared Task on Pronoun Translation. LINDAT/CLARIN digital library at Institute of Formal and Applied Linguistics, Charles University in Prague. http://hdl.handle.net/11372/LRT-1611.

Martin Haspelmath. 1993. More on the typology of inchoative/causative verb alternations. In Bernard Comrie and Maria Polinsky, editors, *Causatives and transitivity*, pages 87–120. John Benjamins, Amsterdam.

Benjamin Heinzerling, Nafise Sadat Moosavi, and Michael Strube. 2017. Revisiting selectional preferences for coreference resolution. In *Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing*, pages 1332–1339, Copenhagen, Denmark. Association for Computational Linguistics.

Otto Jespersen. 1927. *Modern English grammar on historical principles, Part III: Syntax (Second Volume)*. Allen and Unwin, London.

Ekaterina Lapshinova-Koltunski, Christian Hardmeier, and Pauline Krielke. 2018. ParCorFull: a parallel corpus annotated with full coreference. In *Proceedings of 11th Language Resources and Evaluation Conference*, pages 00–00, Miyazaki, Japan. European Language Resources Association (ELRA). To appear.

Timothy Lee, Alex Lutz, and Jinho D. Choi. 2016. QA-It: classifying non-referential it for question answer pairs. In *Proceedings of the ACL 2016 Student Research Workshop*, pages 132–137, Berlin, Germany. Association for Computational Linguistics.

Sharid Loáiciga, Liane Guillou, and Christian Hardmeier. 2017. What is it? disambiguating the different readings of the pronoun "it". In *Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing*, pages 1325–1331, Copenhagen, Denmark. Association for Computational Linguistics.

Christoph Müller. 2007. Resolving *It*, *This*, and *That* in unrestricted multi-party dialog. In *Proceedings of the 45th Annual Meeting of the Association for Computational Linguistics*, ACL07, pages 816–823, Prague, Czech Republic. Association for Computational Linguistics (ACL).

Robert Munro, Steven Bethard, Victor Kuperman, Vicky T. Lai, Robin Melnick, Christopher Potts, Tyler Schnoebelen, and Harry Tily. 2010. Crowdsourcing and language studies: the new generation of linguistic data. In *Proceedings of the NAACL HLT 2010 Workshop on Creating Speech and Language Data with Amazon's Mechanical Turk*, pages 122–130. Association for Computational Linguistics.

Anna Nedoluzhko and Ekaterina Lapshinova-Koltunski. 2016. Abstract coreference in a multilingual perspective: a view on czech and german. In *Proceedings of the Workshop on Coreference Resolution Beyond OntoNotes*, CORBON 2016, pages 47–52, Ann Arbor, Michigan. Association for Computational Linguistics.

Rebecca J. Passonneau. 1989. Getting at discourse referents. In *Proceedings of the 27th Annual Meeting of the Association for Computational Linguistics*, pages 51–59, Vancouver, British Columbia, Canada. Association for Computational Linguistics.

Massimo Poesio. 2015. Linguistic and cognitive evidence about anaphora. In Massimo Poesio, Roland Stuckardt, and Yannick Versley, editors, *Anaphora Resolution: Algorithms, Resources and Application*, pages 23–54. Springer-Verlag, Berlin Heidelberg.

Sameer Pradhan, Alessandro Moschitti, Nianwen Xue, Hwee Tou Ng, Anders Björkelund, Olga Uryupina, Yuchen Zhang, and Zhi Zhong. 2013. Towards robust linguistic analysis using OntoNotes. In *Proceedings of the Seventeenth Conference on Computational Natural Language Learning*, pages 143–152, Sofia, Bulgaria. Association for Computational Linguistics.

R Development Core Team. 2008. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0.

Florian Schäfer. 2009. The causative alternation. *Language and Linguistics Compass*, 3(2):641–681.

Yannick Versley and Anders Björkelund. 2015. Off-the-shelf tools. In Massimo Poesio, Roland Stuckardt, and Yannick Versley, editors, *Anaphora Resolution: Algorithms, Resources and Applications*, pages 237–266. Springer-Verlag, Berlin Heidelberg.

Bonnie Webber. 1986. Findings of the 2016 WMT shared task on cross-lingual pronoun prediction. In *Theoretical Issues in Natural Language Processing*, TINLAP-3, pages 158–163, Las Cruces, New Mexico. Association for Computational Linguistics.

Bonnie L. Webber. 1990. Structure and ostension in the interpretation of discourse deixis. Technical Report MS-CIS-90-58, University of Pennsylvania, Department of Computer and Information Science.