Last Words

What Computational Linguists Can Learn from Psychologists (and Vice Versa)

Emiel Krahmer[*]
Tilburg University

1. Introduction

Sometimes I am amazed by how much the field of computational linguistics has changed in the past 15 to 20 years. In the mid 1990s, I was working at a research institute where language and speech technologists worked in relatively close quarters. Speech technology seemed on the verge of a major breakthrough; this was around the time that Bill Gates was quoted in *Business Week* as saying that speech was not just the future of Windows, but the future of computing itself. At the same time, language technology was, well, nowhere. Bill Gates certainly wasn't championing language technology in those days. And while the possible applications of speech technology seemed endless (who would use a keyboard in 2010, when speech-driven user interfaces would have replaced traditional computers?), the language people were thinking hard about possible applications for their admittedly somewhat immature technologies.

Predicting the future is a tricky thing. No major breakthrough came for speech technology—I am still typing this. However, language technology did change almost beyond recognition. Perhaps one of the main reasons for this has been the explosive growth of the Internet, which helped language technology in two different ways. On the one hand it instigated the development and refinement of techniques needed for searching in document collections of unprecedented size, on the other it resulted in a large increase of freely available text data. Recently, language technology has been particularly successful for tasks where huge amounts of textual data is available to which statistical machine learning techniques can be applied (Halevy, Norvig, and Pereira 2009). As a result of these developments, mainstream computational linguistics is now a successful, application-oriented discipline which is particularly good at extracting information from sequences of words.

But there is more to language than that. For speakers, words are the result of a complex speech production process; for listeners, they are what starts off the similarly complex comprehension process. However, in many current applications no attention is given to the processes by which words are produced nor to the processes by which they can be understood. Language is treated as a product not as a process, in the terminology of Clark (1996). In addition, we use language not only as a vehicle for factual information exchange; speakers may realise all sorts of other intentions with their words: They may want to convince others to do or buy something, they may want to induce a particular emotion in the addressee, and so forth. These days, most of computational linguistics (with a few notable exceptions, more about which subsequently) has little to

---

  [*] Tilburg Centre for Creative Computing (TiCC), Communication and Cognition research group, Tilburg University, The Netherlands. E-mail: e.j.krahmer@uvt.nl.

say about how people produce and comprehend language, nor about what the possible effects of language could be.

It wasn't always like this; early work in computational linguistics took a different (and more ambitious) perspective. Winograd (1980), to give one more or less random example, explicitly treated language understanding and production as cognitive processes, which interacted with other cognitive modules such as visual perception; Hovy (1990), to give another, presented a computational model that generated different texts from the same underlying facts, depending on pragmatic, interpersonal constraints. It is interesting to observe that Winograd and Hovy built on both computational and psychological research, something which is increasingly rare in the field of computational linguistics, a point made convincingly by Reiter (2007). By now, it is generally accepted that the problems that Winograd, Hovy, and others tried to tackle are very complex, and that the current emphasis on more well-delimited problems is probably a good thing. However, it is not difficult to come up with computational applications for which a better understanding would be required of language as a process and the effects language may have on a user (interactive virtual agents which try to persuade a user to do something, for example). To learn more about how speakers and addressees manage to accurately produce and comprehend complex and potentially ambiguous sentences in real time, and how they may use these sentences for a whole range of intentions, we have to turn to psycholinguistics and social psychology, respectively. So let us sample some of the recent findings in these fields, and see if and how they might benefit computational linguistics. Interestingly, we will find many places where more attention to what goes on in computational linguistics would benefit psychologists as well.

2. Language Use and Its Social Impact

Social psychologists study persons and the relations they have with others and with groups. Various social psychologists have concentrated on language (although perhaps not as many as you would expect given the importance of language for social interactions). A number of different approaches can be discerned, one of which concentrates on the psychological functions of function words (Chung and Pennebaker 2007). Function words are understood here to include pronouns, prepositions, articles, conjunctions, and auxiliary verbs.

2.1 On the Psychological Functions of Pronouns

One reliable finding of this perspective is that first person singular pronouns are associated with negative affective states. For example, in one study it was found that currently depressed students used *I* and *me* more often than students who were not currently depressed, and of the latter group those who had known periods of depression used them more frequently than those who had never had such an episode (Rude, Gortner, and Pennebaker 2004). Another study found that suicidal poets used first person singular pronouns in their poems more frequently than non-suicidal poets (Stirman and Pennebaker 2001). Of course, whether a speaker tends to use *I* or *we* more frequently is also indicative of self- versus other-centeredness. An analysis of blogs following the events of September 11 revealed that bloggers' use of *I* and *me* dropped in the hours following the attack, while simultaneously their use of *we* and *us* increased (Cohn, Mehl, and Pennebaker 2004); this switch is interpreted by the authors as indicating that people

were focusing less on themselves during this period, but instead focusing more on their friends and families. In a completely different study of adult speakers (both male and female) who underwent testosterone therapy, it was found that as testosterone levels dropped, so did their use of *I* pronouns, while simultaneously the use of non-*I* pronouns increased (Pennebaker et al. 2004).

Pennebaker and colleagues report comparable effects of age, gender, status, and culture on personal pronoun use (Chung and Pennebaker 2007). Their corpus (or "archive", as they call it) contains over 400,000 text files, from many different authors and collected over many years. It is interesting to observe that Pennebaker was an early adapter of computers for his analyses, simply because performing them manually was too time-consuming. The general approach in these analyses is to determine beforehand what the "interesting" words are and then simply to count them in the relevant texts, without taking the linguistic context into account. This obviously creates errors: The relative frequency of first-person pronouns may be indicative of depression, as we have just seen, but a sentence such as *I love life* seems a somewhat implausible cue for a depressed state of mind. Chung and Pennebaker (2007, page 345) themselves give the example of *mad*, which is counted as an anger and negative emotion word, and they point out that this is wrong for *I'm mad about my lover*. Clearly, standard methods from computational linguistics could be used to address this problem, for instance by looking at words in context and *n*-grams. Another problem which Chung and Pennebaker mention, and which will be familiar to many computational linguists, is the problem of deciding which are the interesting words to count. Here techniques such as feature construction and selection could be of help. As I will argue in what follows, the observations of Pennebaker and colleagues are potentially interesting for computational linguistics as well, but let us first look at another relevant set of psychological findings.

2.2 On the Psychological Functions of Interpersonal Language

A different strand of research on language and social psychology focuses on inter-personal verbs (a subset of what computational linguists more commonly refer to as transitive verbs): verbs which express relations between people (Semin 2009). In their model of interpersonal language (the Linguistic Categorization Model), Semin and Fiedler (1988) make a distinction between different kinds of verbs and their position on the concrete–abstract dimension. Descriptive action verbs (*Romeo kisses Juliet*) are assumed to be the most concrete, since they refer to a single, observable event. This is different for state verbs (*Romeo loves Juliet*), which describe psychological states instead of single perceptual events, and are therefore more abstract. Most abstract, according to Semin and Fiedler, are adjectives (*Romeo is romantic*), because they generalize over specific events and objects and only refer to characteristics of the subject.

The thing to note is that the same event can, in principle, be referred to in all these different forms; a speaker has the choice of using a more concrete or a more abstract way to refer to an event (e.g., John can be described as, from more to less concrete, hitting a person, hating a person, or being aggressive). Interestingly, the abstractness level a speaker opts for tells us something about that speaker. This has been found, for instance, in the communication of ingroup (think of people with the same cultural identity or supporting the same soccer team) and outgroup (different identity, different team) behavior. There is considerable evidence that speakers describe negative ingroup and positive outgroup behavior in more concrete terms (e.g., using action verbs), thereby indicating that the behavior is more incidental, whereas positive ingroup

and negative outgroup behaviors are described in relatively abstract ways (e.g., more frequently using adjectives), suggesting a more enduring characteristic (see, e.g., Maass et al. 1989). Maass and colleagues showed this phenomenon, which they dubbed the Linguistic Intergroup Bias, for different *Contrada* (neighborhoods) participating in the famous *Palio di Siena* horse races, reporting about their own performance and that of the other neighborhoods. Moreover, Wigboldus, Semin, and Spears (2000) have shown that addressees do indeed pick up these implications, and Douglas and Sutton (2006) reveal that speakers who describe the behavior of others in relatively abstract terms are perceived to have biased attitudes and motives as opposed to speakers who describe this behavior in more concrete ways.

It has been argued that concrete versus abstract language is not only relevant for, for example, the communication of stereotypes, but also has more fundamental effects, for instance influencing the way people perceive the world (Stapel and Semin 2007). In a typical experiment, Stapel and Semin first subtly prime participants with either abstract or concrete language. This can be done using scrambled sentences, where participants are given four words (*romantic is lamp Romeo*) with the instruction to form a grammatical sentence from three of them, or by giving participants a word-search puzzle where the words to search for are the primes. After this, participants perform a seemingly unrelated task, where their perceptual focus (either on the global picture or on the details) is measured. Stapel and Semin show that processing abstract language (adjectives) results in a more global perception, whereas processing concrete language (descriptive action verbs) leads to more perceptual attention to details.

At this point, a computational linguist (and probably other linguists as well) might start to wonder about the comparison between verbs and adjectives, and by the claim that adjectives are abstract. What about adjectives like *blonde*, *young*, and *thin*? These seem to be much more concrete than adjectives such as *aggressive* or *honest*. And what about nouns? There a distinction between concrete (*office chair*) and abstract (*hypothesis*) seems to exist as well. This raises the question whether it is the differences in interpersonal language use or the more general distinction between concrete and abstract language which causes the observed effects on perception; a recent series of experiments suggests it is the latter (Krahmer and Stapel 2009).

2.3 What Can Computational Linguists Learn?

The social psychological findings briefly described here could have an impact on computational linguistics, with potential applications for both text understanding and generation. So far, it seems fair to say that most computational linguists have concentrated so much on trying to understand text or on generating coherent texts that the subtle effects that language may have on the reader were virtually ignored. Function words were originally not the words computational linguists found most interesting. They were considered too frequent; early Information Retrieval applications listed function words on "stop lists"—lists of words that should be ignored during processing—and many IR applications still do. The work of Pennebaker and colleagues indicates that pronouns (as well as other function words) do carry potentially relevant information, for instance about the mental state of the author of a document. Interestingly, for computational applications such as opinion mining and sentiment analysis (Pang and Lee 2008) as well as author attribution and stylometry (Holmes 1998), function words have been argued to be relevant as well, but it seems that research on the social psychology of language has made little or no impact on this field.

Consider sentiment analysis, for instance, which is the automatic extraction of "opinion-oriented" information (e.g., whether an author feels positive or negative about a certain product) from text. This is a prime example of an emerging research area in computational linguistics which moves beyond factual information exchange (although the preferred approach to this problem very much fits with the paradigm sketched by Halevy et al. [2009]: take a large set of data and apply machine learning to it). Pang and Lee (2008) offer an extensive overview of research related to sentiment analysis, but do not discuss any of the psychological studies mentioned herein (in fact, of the 332 papers they cite, only one or two could conceivably be interpreted as psychological in the broadest interpretation). What is especially interesting is that their discussion of why sentiment analysis is difficult echoes the discussion of Chung and Pennebaker (2007) on the problems of counting words (by sheer coincidence they even discuss essentially the same example: *madden*).

These findings may also have ramifications for the generation of documents. If you develop an application which automatically produces texts from non-textual data, you might want to avoid excessive use of the first-person pronoun, lest your readers think your computer is feeling down. If you want your readers to skim over the details of what is proposed in a generated text, use abstract language. In addition, you may want to use action verbs when describing your own accomplishments, and adjectives to refer to those of others (but do it in a subtle way, because people might notice).

3. Language Comprehension and Production

While the link between computational linguistics and social psychology has seldom been explored, there has been somewhat more interaction with psycholinguistics. Perhaps most of this interaction has involved natural language understanding. Various early parsing algorithms were inspired by human sentence processing, which is hardly surprising: human listeners are remarkably efficient in processing and adequately responding to potentially highly ambiguous sentences. Later, when large data sets of parsed sentences became available, the focus in computational linguistics shifted to developing statistical models of language processing. Interestingly, recent psycholinguistic sentence processing models are inspired in turn by statistical techniques from computational linguistics (Chater and Manning 2006; Crocker in press; Jurafsky 2003; Pado, Crocker, and Keller 2009).

3.1 On Producing Referring Expressions

The situation is somewhat different for natural language generation, although superficially the same kind of interaction can be observed here (albeit with a few years delay). The seminal work by Dale and Reiter (1995) on the generation of referring expressions was explicitly inspired by psycholinguistic work. Dale and Reiter concentrated on the generation of distinguishing descriptions, such as *the large black dog*, which single out one target object by ruling out the distractors (typically a set of other domestic animals of different sizes and colors). Given that the number of distractors may be quite large and given that each target can be referred to in multiple ways, one of the main issues in this area is how to keep the search manageable. Current algorithms for referring expression generation, building on the foundations laid by Dale and Reiter, are good at quickly computing which set of properties uniquely characterizes a target among a

set of distractors. Some of these algorithms are capable of generating distinguishing descriptions that human judges find more helpful and better formulated than human-produced distinguishing descriptions for the same targets (Gatt, Belz, and Kow 2009).

To some this might suggest that the problem is solved. This conclusion, however, would be too hasty. Most of the algorithms use some very unrealistic assumptions which limit their applicability. Interestingly, these assumptions can be traced back directly to classic psycholinguistic work on the production of referring expressions (Olson 1970). Clark and Bangerter (2004) criticize a number of the unstated assumptions in Olson's approach: Reference is treated as a one-step process (a speaker plans and produces a complete description, and nothing else, in one go) and during that process the speaker does not take the prior interaction with the addressee into account. By merely substituting *computer* for *speaker* these comments are directly applicable to most current generation algorithms as well.

The problem, unfortunately, is that recent psycholinguistic research suggests that these assumptions are wrong. Often this research looks at how speakers produce referring expressions while interacting with an addressee, and one thing that is often found is that speakers adapt to their conversational partners while producing referring expressions (Clark and Wilkes-Gibbs 1986; Brennan and Clark 1996; Metzing and Brennan 2003). This kind of "entrainment" or "alignment" (Pickering and Garrod 2004) may apply at the level of lexical choice; if a speaker refers to a couch using the word *sofa* instead of the more common *couch*, the addressee is more likely to use *sofa* instead of *couch* as well later on in the dialogue (Branigan et al. in press). But the speaker and addressee may also form a general "conceptual pact" on how to refer to some object, deciding together, for instance, to refer to a tangram figure as *the tall ice skater*.

Although adaptation itself is uncontroversial, psycholinguists argue about the extent to which speakers are capable of taking the perspective of the addressee into account (Kronmüller and Barr 2007; Brennan and Hanna 2009; Brown-Schmidt 2009), with some researchers presenting evidence that speakers may have considerable difficulty doing this (Horton and Keysar 1996; Keysar, Lin, and Barr 2003). In Wardlow Lane et al. (2006) people are instructed to refer to simple targets (geometrical figures that may be small or larger) in the context of three distractor objects, two of which are visible to both speaker and addressee (shared) whereas the other is visible to the speaker only (privileged). If speakers would be able to take the addressees' perspective into account when referring, the privileged distractor should not play a role in determining which properties to include in the distinguishing description. However, Wardlow Lane and colleagues found that speakers do regularly take the privileged distractor into account (for instance adding a modifier *small* when referring to the target, even though all the shared objects are small and only the privileged one is large). Interestingly, speakers do this more often when explicitly told that they should *not* leak information about the privileged object, which Wardlow Lane et al. interpret as an ironic processing effect of the kind observed by Dostoevsky ("Try to pose for yourself this task: not to think of a polar bear, and you will see that the cursed thing will come to mind every minute").

Another interesting psycholinguistic finding is that speakers often include more information in their referring expressions than is strictly needed for identification (Arts 2004; Engelhardt, Bailey, and Ferreira 2006), for instance referring to a dog as *the large black curly haired dog* in a situation where there is only one large black dog. Again, that speakers are not always "Gricean" ("be as informative as required, but not more informative") is generally agreed upon, but there is an ongoing debate about why and how speakers overspecify, some arguing that it simplifies the search of the speaker

(Engelhardt, Bailey, and Ferreira 2006) whereas others suggest that overspecified references are particularly beneficial for the addressee (Paraboni, van Deemter, and Masthoff 2007).

3.2 What Can Computational Linguists Learn?

Why are these psycholinguistic findings about the way human speakers refer relevant for generation algorithms? First of all, human-likeness is an important evaluation criterion, so algorithms that are good at emulating human referring expressions are likely to outperform algorithms that are not. Moreover, it is interesting to observe that generating overspecified expressions is computationally cheaper than producing minimal ones (Dale and Reiter 1995). In a similar vein, it can be argued that alignment and adaptation may reduce the search space of the generation algorithm, because they limit the number of possibilities that have to be considered.

It is worth emphasizing that psycholinguistic theories have little to say about how speakers quickly decide which properties, from the large set of potential ones, to use in a referring expression. In addition, whereas notions such as adaptation, alignment, and overspecification are intuitively appealing, it has turned out to be remarkably difficult to specify how these processes operate exactly. In fact, a common criticism is that they would greatly benefit from "explicit computational modeling" (Brown-Schmidt and Tanenhaus 2004). Of course, solving choice problems and computational modeling are precisely what computational linguistics has to offer. So although generation algorithms may benefit a lot from incorporating insights from psycholinguistics, they in turn have the potential to further research in psycholinguistics as well.

4. Discussion

In this brief, highly selective, and somewhat biased overview of work on language in several areas of psychology, we have seen that words may give valuable information about the person who produces them (but how to select and count them is tricky), that abstract or concrete language may tell you something about the opinions and attitudes a speaker has and may even influence how you perceive things (but the linguistic intuitions about what is abstract, and what concrete, need some work), and that speakers are remarkably efficient when producing referring expressions, in part because they adapt to their addressee and do not necessarily try to be as brief as possible (but making these intuitive notions precise is difficult). Psychological findings such as these are not merely intriguing, but could be of real use for computational linguistic applications related to document understanding or generation (and, conversely, techniques and insights from computational linguistics could be helpful for psychologists as well). Of course, some computational linguists do extensively rely on psychological findings for building their applications (you know who you are), just as some psychologists use sophisticated computational and statistical models rather than human participants for their studies (this is especially true in psycholinguistics). But these are exceptions, and certainly do not belong to mainstream computational linguistics or psychology. Which raises one obvious question: Why isn't there more interaction between these two communities?

There seem to be at least three reasons for this. First, and most obvious, many researchers are not aware of what happens outside their own specialized field. The articles in psychology are fairly accessible (usually no complex statistical models or overformalized algorithms there), but many computational linguists may feel that it

would be a better investment of their limited time to read some more of the 17,000 (and counting) journal, conference, and workshop papers they have not yet read in the invaluable ACL Anthology. For psychologists presumably similar considerations apply, with the additional complication that many of the anthology papers require a substantial amount of technical prior knowledge. In addition, it might be that the different publication cultures are a limiting factor here as well: for psychologists, journals are the main publication outlet; for them most non-journal publications have a low status and hence might be perceived as not worth exploring.

Another perhaps more interesting reason is that psychologists and computational linguists have subtly different general objectives. Psychologists want to get a better understanding of people; how their social context determines their language behavior, how they produce and comprehend sentences, and so on. Their models are evaluated in terms of whether there is statistical evidence for their predictions in actual human behavior. Computational linguists evaluate their models ("algorithms") on large collections of human-produced *data*; one model is better than another if it accounts for more of the data. Of course, a model can perform well when evaluated on human data, but be completely unrealistic from a psychological point of view. If a computational linguist develops a referring expression generation algorithm (or a machine translation system or an automatic summarizer) which accounts for the data in a way which is psychologically unrealistic, the work will generally not be of interest to psychologists. Conversely, if psychological insights are difficult to formalize or require complex algorithms or data structures, computational linguists are likely not to be enthusiastic about applying them. Obviously, this hinders cross-pollination of ideas as well.

Third, and somewhat related to the previous point, it sometimes seems that computational linguists see trees where psychologists see a forest. Psychologists appear to be most interested in showing a general effect (and are particularly appreciative of clever and elegant experimental designs which reveal these effects); if merely counting words already gives you a statistically reliable effect, then why bother with a more complicated way of counting *n*-grams and worrying about back-off smoothing to deal with data sparseness? Doing so would conceivably give you a better estimate of the significance and size of your effect, but would probably not change your story in any fundamental way. Computational linguists, by contrast, evaluate their models on (often shared) datasets (and tend to be more impressed by technical prowess—e.g., new statistical machine learning models—or by smart ways of automatically collecting large quantities of data); each data point that is processed incorrectly by their model offers a potential advantage for someone else's model.

In view of observations such as these, it is perhaps not surprising that computational linguists and psychologists have remained largely unaware of each other's work so far. Predicting the future is a tricky thing, but it seems not unlikely that most computational linguists and psychologists will continue going their own way in the future. Nevertheless, I hope to have shown here that both communities could benefit from the occasional foray into the others' territory. For psychologists, the tools and techniques developed by computational linguists could further their research, by helping to make their models and theories more explicit and hence easier to test and compare. For computational linguists, insights from both the social psychology of language and from psycholinguists could contribute to a range of applications, from opinion mining to text understanding and generation. Obviously, this contribution could be on the level of "words", but a more substantial contribution is conceivable as well. As we have seen, psychologists are particularly strong in explanatory theories (on affect, on interaction, etc.) and perhaps taking these as starting points for our applications (e.g.,

on affective and interactive generation) could make them theoretically more interesting and empirically more adequate.

References

Arts, Anja. 2004. *Overspecification in Instructive Texts*. Unpublished Ph.D. thesis, Tilburg University.

Branigan, Holly P., Martin J. Pickering, Jamie Pearson, and Janet F. McLean. (in press). Linguistic alignment between humans and computers. *Journal of Pragmatics*.

Brennan, Susan and Herbert H. Clark. 1996. Conceptual pacts and lexical choice in conversation. *Journal of Experimental Psychology*, 22(6):1482–1493.

Brennan, Susan E. and Joy E. Hanna. 2009. Partner-specific adaptation in dialog. *Topics in Cognitive Science*, 1(2):274–291.

Brown-Schmidt, S. and M. Tanenhaus. 2004. Priming and alignment: Mechanism or consequence? commentary on Pickering and Garrod 2004. *Behavioral and Brain Sciences*, 27:193–194.

Brown-Schmidt, Sarah. 2009. Partner-specific interpretation of maintained referential precedents during interactive dialog. *Journal of Memory and Language*, 61:171–190.

Chater, Nick and Christopher D. Manning. 2006. Probabilistic models of language processing and acquisition. *Trends in Cognitive Science*, 10:335–344.

Chung, C. K. and James W. Pennebaker. 2007. The psychological function of function words. In K. Fiedler, editor, *Social Communication: Frontiers of Social Psychology*. Psychology Press, New York, pages 343–359.

Clark, Herbert H. 1996. *Using Language*. Cambridge University Press, Cambridge, UK.

Clark, Herbert H. and Adrian Bangerter. 2004. Changing ideas about reference. In Ira A. Noveck and Dan Sperber, editors, *Experimental Pragmatics*. Palgrave Macmillan, Basingstoke, pages 25–49.

Clark, Herbert H. and Deanna Wilkes-Gibbs. 1986. Referring as a collaborative process. *Cognition*, 22:1–39.

Cohn, M., M. Mehl, and James W. Pennebaker. 2004. Linguistic markers of psychological change surrounding September 11, 2001. *Psychological Science*, 15:687–693.

Crocker, Matthew W. (in press). Computational psycholinguistics. In Alex Clark, Chris Fox, and Shalom Lappin, editors, *Computational Linguistics and Natural Language Processing Handbook*. Wiley Blackwell, London, UK.

Dale, Robert and Ehud Reiter. 1995. Computational interpretations of the Gricean maxims in the generation of referring expressions. *Cognitive Science*, 18:233–263.

Douglas, Karen and Robbie Sutton. 2006. When what you say about others says something about you: Language abstraction and inferences about describers' attitudes and goals. *Journal of Experimental Social Psychology*, 42:500–508.

Engelhardt, Paul E., Karl G. D. Bailey, and Fernanda Ferreira. 2006. Do speakers and listeners observe the Gricean Maxim of Quantity? *Journal of Memory and Language*, 54:554–573.

Gatt, Albert, Anja Belz, and Eric Kow. 2009. The tuna-reg challenge 2009: Overview and evaluation results. In *Proceedings of the 12th European Workshop on Natural Language Generation (ENLG)*, pages 174–182, Athens.

Halevy, Alon, Peter Norvig, and Fernando Pereira. 2009. The unreasonable effectiveness of data. *IEEE Intelligent Systems*, 24:8–12.

Holmes, David I. 1998. The evolution of stylometry in humanities scholarship. *Literary and Linguistic Computing*, 13:111–117.

Horton, W. S. and B. Keysar. 1996. When do speakers take into account common ground? *Cognition*, 59:91–117.

Hovy, Eduard H. 1990. Pragmatics and natural language generation. *Artificial Intelligence*, 43:153–197.

Jurafsky, Dan. 2003. Probabilistic modeling in psycholinguistics: Linguistic comprehension and production. In Rens Bod, Jennifer Hay, and Stefanie Jannedy, editors, *Probabilistic Linguistics*. MIT Press, Cambridge, MA, pages 39–96.

Keysar, B., S. Lin, and D. J. Barr. 2003. Limits on theory of mind use in adults. *Cognition*, 89:25–41.

Krahmer, Emiel and Diederik Stapel. 2009. Abstract language, global perception: How language shapes what we see. In *Proceedings of the Annual Meeting of the Cognitive Science Society*, pages 286–291, Amsterdam.

Kronmüller, E. and Dale Barr. 2007. Perspective-free pragmatics: Broken precedents and the recovery-from-preemption hypothesis. *Journal of Memory and Language*, 56:436–455.

Maass, A., D. Salvi, L. Arcuri, and G. Semin. 1989. Language use in intergroup contexts: The linguistic intergroup bias. *Journal of Personality and Social Psychology*, 57:981–993.

Metzing, Charles A. and Susan E. Brennan. 2003. When conceptual pacts are broken: Partner effects on the comprehension of referring expressions. *Journal of Memory and Language*, 49:201–213.

Olson, David R. 1970. Language and thought: Aspects of a cognitive theory of semantics. *Psychological Review*, 77:257–273.

Pado, Ulrike, Matthew W. Crocker, and Frank Keller. 2009. A probabilistic model of semantic plausibility in sentence processing. *Cognitive Science*, 33:794–838.

Pang, B. and L. Lee. 2008. Opinion mining and sentiment analysis. *Foundations and Trends in Information Retrieval*, 2:1–135.

Paraboni, Ivandré, Kees van Deemter, and Judith Masthoff. 2007. Generating referring expressions: Making referents easy to identity. *Computational Linguistics*, 33:229–254.

Pennebaker, James W., C. Groom, D. Loew, and J. Dabbs. 2004. Testosterone as a social inhibitor: Two case studies of the effect of testosterone treatment on language. *Journal of Abnormal Psychology*, 113:172–175.

Pickering, Martin and Simon Garrod. 2004. Towards a mechanistic psychology of dialogue. *Behavioural and Brain Sciences*, 27:169–226.

Reiter, Ehud. 2007. The shrinking horizons of computational linguistics. *Computational Linguistics*, 33:283–287.

Rude, S., E. Gortner, and James W. Pennebaker. 2004. Language use of depressed and depression-vulnerable college students. *Cognition and Emotion*, 18:1121–1133.

Semin, Gün. 2009. Language and social cognition. In F. Strack and J. Förster, editors, *Social Cognition: The Basis of Human Interaction*. Psychology Press, London, pages 269–290.

Semin, Gün and Klaus Fiedler. 1988. The cognitive functions of linguistic categories in describing persons: Social cognition and language. *Journal of Personality and Social Psychology*, 34:558–568.

Stapel, Diederik and Gün Semin. 2007. The magic spell of language. *Journal of Personality and Social Psychology*, 93:23–33.

Stirman, Shannon and James W. Pennebaker. 2001. Word use in the poetry of suicidal and nonsuicidal poets. *Psychosomatic Medicine*, 63:517–522.

Wardlow Lane, Liane, Michelle Groisman, and Victor S. Ferreira. 2006. Don't talk about pink elephants! : Speakers' control over leaking private information during language production. *Psychological Science*, 17:273–277.

Wigboldus, Daniël, Gün Semin, and Russell Spears. 2000. How do we communicate stereotypes? linguistic bases and inferential consequences. *Journal of Personality and Social Psychology*, 78:5–18.

Winograd, Terry. 1980. What does it mean to understand language? *Cognitive Science*, 4:209–241.