Influence of perspective and goals on reference production in conversation

Si On Yoon (syoon@illinois.edu)
Department of Psychology, University of Illinois
Champaign, IL 61820 USA

Sungryong Koh (koh@snu.ac.kr)
Department of Psychology, Seoul National University
Seoul, Korea

Sarah Brown-Schmidt (brownsch@illinois.edu)
Department of Psychology, University of Illinois
Champaign, IL 61820 USA

Abstract

We examined the extent to which speakers take into consideration the addressee's perspective in language production. Previous research on this process has revealed clear deficits (Horton & Keysar, 1996; Wardlow Lane & Ferreira, 2008). Here we examined the use of perspective and evaluated a new hypothesis—that the relevance of the addressee's perspective during language production depends on speaker goals. In two experiments, Korean speakers described a target object in situations where the perspective-status of a contrasting object (e.g., a large plate when describing a smaller plate) was manipulated. Experiment 1 examined whether speakers would use scalar-modified expressions even when the contrast was hidden from the addressee. The results demonstrated that information from both the speaker's and addressee's perspective influences production. Experiment 2 examined whether utterance goals would modulate this process. The results indicated that when a speaker makes a request, the addressee's perspective has a stronger influence on utterance form than when the speaker informs the addressee. These results suggest that a speaker's privileged knowledge shapes language use, but that crucially, the degree to which the addressee's perspective is considered is shaped by the relevance of the addressee's perspective to utterance goals.

Keywords: language production; referential expression; utterance goal; perspective.

Introduction

A central question in language production is the degree to which speakers take into account the addressee's perspective. For example, imagine a situation in which two interlocutors face each other, and a speaker refers to a mutually known object (e.g., a plate), in a context with a size-contrasting object that is visible to the speaker, but hidden from the addressee (e.g., a smaller plate). In this situation, the speaker can refer to her intended referent as either *the plate* or *the big plate,* depending on whether she designs her expression from her own egocentric perspective, or if she considers the addressee's perspective. The egocentric expression, *the big plate*, would uniquely identify the referent from the speaker's privileged perspective but would be over-modified from the addressee's perspective. Instead, if the speaker considers what information is mutually known, e.g., the *common ground*, she would likely refer the object using the expression, *the plate*, which is appropriately modified from the addressee's perspective. The present research examines whether speakers, in situations such as this one, design expressions with respect to common ground. We examine this question in a new participant population—Koreans—and examine the potential role for previously unconsidered modulating factors, in particular, utterance goals.

According to Grice's Maxims of Quantity, speakers are assumed to produce expressions that are as informative as required, and that are therefore neither under- nor over-informative. In scenario above, a cooperative speaker is expected to use a bare noun, since only a single plate is in common ground. Use of a scalar modifier in this case would be non-cooperative--a violation of the Maxim of Quantity. However, in natural conversation, speakers are not strictly Gricean and not always likely to obey the Maxim of Quantity. While over-informative expressions may result in confusion in some circumstances (Wardlow Lane & Ferreira, 2008), in many cases they do not, and sometimes may actually play an advantageous role (Deutsch, 1976).

Perspective in Production: Two Views

When do speakers take into consideration the addressee's perspective when designing referential expressions? Two prominent views have emerged. According to the Constraint-Based Processing view, multiple sources of information, including contextual information and the perspective of the addressee, combine to constrain language comprehension and production processes (Tanenhaus & Trueswell, 1995; Hanna, Tanenhaus & Trueswell, 2003; Horton, 2007). On this view, representations of the addressee's perspective combine with other sources of information, such as the speaker's privileged perspective, to constrain production. This view predicts early, and strong contributions of common ground to referential production,

in cases where this information is well-established and relevant (Brown-Schmidt & Hanna, 2011).

This view is in contrast to the Egocentric Heuristics view, which proposes that expressions are initially designed from the speaker's egocentric perspective (Horton & Keysar, 1996). On this view, the speaker monitors the addressee for understanding, and in cases of confusion, can adjust her expression to meet the addressee's needs. This view is related to egocentric-first processing theories of language comprehension (e.g., Keysar, Lin, & Barr, 2003), and is motivated by an assumption that incorporating common ground into routine language processes would be too resource intensive (Keysar, Barr, Balin & Paek, 1998).

Perspective in Production: Previous Research

Despite claims that common ground is central to language (Clark, 1996), and a clear division between processing theories which hypothesize a central role for common ground in language production (Heller, Gorman, & Tanenhaus, in press), and those that do not (Horton & Keysar, 1996), relatively little work has investigated this question experimentally. Further, most (if not all) of this work has been conducted in a single participant population—a limitation which will soon become clear.

Some research suggests the addressee's perspective is only relevant during a delayed, second stage of production. Horton and Keysar (1996) asked speakers to describe target objects to addressees in the context of a second, contrasting object that was either in common ground or the speaker's privileged ground (e.g., a small circle in the context of a larger circle). In one condition, speakers were placed under time pressure. If utterances are initially designed with respect to common ground, it was argued that speakers should design their utterances with respect to the addressee's perspective, even when under time pressure. In the absence of time pressure, speakers were significantly more likely to include a contextually relevant adjective (e.g., *small*) when the context object was common ground. This effect was eliminated in the speeded condition, consistent with the claim that use of perspective is time-consuming.

Similar results were obtained by Wardlow Lane, Groisman & Ferreira (2008) who evaluated whether speakers could control privileged information when it would be communicatively disadvantageous. Speakers referred to a mutually known object in contexts that contained size-contrasting objects in the speaker's privileged ground. In one condition, speakers were instructed to conceal the identity of the privileged object. If use of perspective is under speakers' control, speakers should be less likely to use a modified expression when instructed to conceal the identity of the privileged object, because scalar modification implies contrast (Sedivy, 1999). Surprisingly, scalar modification rates were *higher* when speakers were instructed to conceal, suggesting that privileged information is automatically incorporated into the referring process, with little ability of the speaker to modulate this process.

Other results show moderate effects of perspective on production. Nadig & Sedivy (2002) found that even 5 to 6 year old children, who are often considered egocentric (Epley, Morewedge, & Keysar, 2004), in fact showed successful use of common ground in both production and comprehension. Similarly, Heller, et al. (in press) found that while speakers sometimes referred to privileged information, they did show some sensitivity to common ground, and that when privileged information was referenced, it may have been used as a pedagogical tool.

Utterance Goals The speaker's communicative goals in producing an utterance are an essential (if not *the* essential) component of language production, however little to no research has examined how goals affect the use of perspective in production. Utterance goals include the expression of a statement, question, command or desire (Searle, 1969). Crucially—perspective plays a different role in each. For example, during language *comprehension*, privileged ground is particularly relevant to interpretation of questions (Brown-Schmidt, et al., 2008), whereas common ground is relevant for imperatives (Hanna, et al., 2003). Here we examined whether goals influence perspective taking in language *production*. Specifically, we asked whether the goals of informing vs. requesting make different demands on perspective-taking. To inform is to state a situation. Presumably then, informing can effectively proceed from the speaker's perspective, thus allowing the speaker to bring new, previously-privileged information into common ground. By contrast, to request is to command. If the speaker wants the addressee to comply, it is critical the addressee can understand the request from her own perspective. Thus, here we test the hypothesis that speakers should consider the addressee's perspective *more* when requesting than informing.

The present research

The experiments presented here examine whether common ground plays a role in language production in a previously unstudied group—Korean speakers (Exp. 1), and explores factors that might modulate this process (Exp. 2). Speakers described a target object in contexts where the perspective status of a contrast object was manipulated. Exp. 1 examined whether speakers would use scalar-modified noun phrases (NPs) even when the modifier was not informative to the addressee. This experiment examines this question in a previously unstudied group, as well as vets a paradigm for use in Exp. 2. In Exp. 2, we examined whether speaker's sensitivity to perspective is influenced by utterance goals, specifically to inform vs. request.

Experiment 1

We examined whether speakers would use scalar modified NPs to identify objects in contexts where modifiers are not informative to the addressee. Utterances specify an intended referent relative to a set of alternatives from which it must be differentiated (Olson, 1970). Based on previous findings

in English (Horton & Keysar, 1996), we expected speakers would produce scalars more often when modification would be informative to the addressee, and that there would be partial interference from the speaker's privileged perspective when it conflicted with that of the addressee.

Method

Participants Eighteen undergraduates at Seoul National University participated in the experiment as speakers in return for course credit. All were native Korean speakers.

Materials 78 objects were grouped into 6 sets of 13 objects each, with three critical pairs of objects per set (i.e., 3 targets and 3 competitors), plus 7 distractor objects. Subjects completed 18 critical and 18 additional filler trials.

Common ground and privileged ground were manipulated in three conditions, which manipulated the number of competitor objects and whether they were mutually visible to speaker and addressee. In the first condition, there was a single target object (e.g. a cup in Fig.1, henceforth "1 object"); this condition established a baseline. In the second condition, there were two *mutually visible* critical objects, a target and a competitor (e.g. a small candle and a big candle in Fig. 1, henceforth "2 objects"). The competitor object was always of the same type, but bigger than the target. In the third condition, there was a *shared* target object and a *privileged* competitor (e.g. a bigger plate and a small plate in Fig. 1, henceforth "2 objects-privileged"). While the speaker saw both target and competitor, a curtain occluded the competitor from the addressee's view. Items were counterbalanced and randomly displayed.

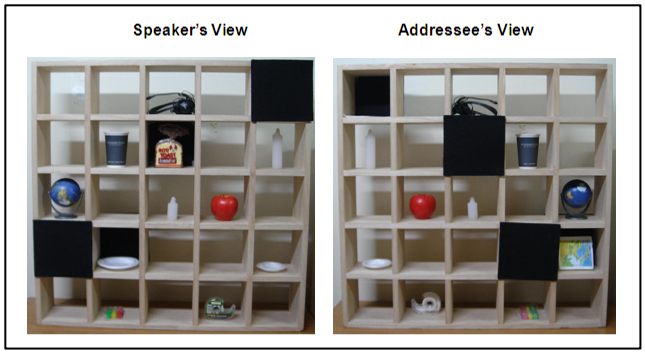

Figure 1. Example of the experimental setup. Critical objects included one cup, two candles and two plates; the bigger plate was hidden from the addressee.

Procedure The participant and confederate sat on either side of a 5X5 grid with objects in some slots. Two 10-inch mini notebook computers, one assigned to a participant and one to the confederate, displayed instruction pictures during the task. Four slots were blocked with a curtain on one side of the grid. Critically, in the 2 objects-Privileged condition, the curtain blocked the contrast from the addressee's view.

The task was a referential communication task. At the beginning of each set, the participant was asked to place the objects in order to duplicate a scene presented on her laptop. While the participant was putting the objects in the grid, the confederate turned his back so he could not see the objects. While the confederate knew which objects were in the grid, the participant believed that the confederate did not know this information. After the participant placed each object, the experimenter blocked four slots with black curtains.

The confederate was then asked to face the subject, and the experimenter asked the confederate to name each object in the grid that they could see. This was to make sure that the participants fully understood they had different perspectives, and to establish which objects were common ground.

After the confederate named and identified all of the objects in the grid, the confederate and participant took turns asking each other to pick an object. One second following a beep sound, instruction pictures were displayed on the computer screen. When one person asked the other to pick an object, the other partner picked the object and put it back in its original place. Participants were allowed to say anything they needed to get the task done. The conversation was recorded on a portable voice recorder. After the task was completed, participants were debriefed on the purpose of the experiment.

Coding From the recordings, the speakers' words were transcribed, and target object descriptions were categorized into three kinds: Bare names (e.g. "*Pick up the plate, please*."), scalar-modified expressions (e.g., "*Pick up the small plate, please*."), and errors in which the participant asked her partner to pick up the wrong object. The sentence structure in Korean is SOV (Subject-Object-Verb), so the first word of utterance was highly likely to be the name of the object. As in English, size adjectives like "*small*" usually come before the noun.

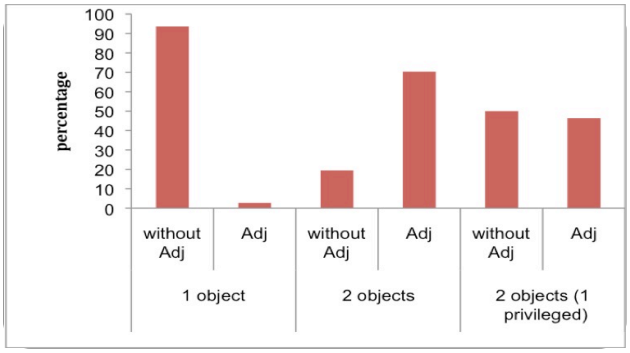

Figure 2. Target descriptions in Experiment 1.

Results and Discussion

For each condition, the percentage of targets described with and without modifiers were computed (Figure 2).

In the conditions where partners had common ground, participants were highly sensitive to referential context. Speakers typically produced a bare noun when there was only one object (e.g., *plate*, 93.52%), but produced a scalar-modified NP when there were two different-sized objects of

the same kind in common ground (e.g., *small plate*, 73.15%). In contrast, in the privileged ground condition, participants used a bare noun (50.00%) or a scalar-modified expression (46.30%) roughly equally. In these situations, the addressee could not see the contrast object, thus the egocentric response was to produce a scalar adjective, whereas the perspective-appropriate response was to produce a bare noun. Therefore, a similar percentage of participants adopted the egocentric perspective as their partner's perspective. This result suggests that both common ground and privileged ground influence utterance design.

A mixed-effects model (see Baayen, Davidson, & Bates, 2008) on the use of size adjectives in the critical utterances confirmed these observations. There were significant differences between the 1-object and 2-objects (b =0.787, SE = 0.0506, $p < .01$), and also 2-objects and the 2-privileged conditions (b =-0.336, SE = 0.0583, $p < .01$). While speakers produced extraneous adjectives ~46% of the time in the 2-privileged condition, this was a significantly lower rate of adjective use than in the 2-objects condition where the competitor was common ground.

These results show that both common and privileged information play a role in planning. This experiment replicates previous findings of moderate audience design in a new participant population and language (Korean), and set the stage for our investigation of the role of speaker goals in the audience design process in Exp. 2. Further, the results of Exp. 1 suggest that in language production, interference from one's own privileged knowledge may reflect a standard, rather than a culturally determined (e.g., Wu & Keysar, 2007), component of language production. In Exp. 2 we explore whether privileged interference is modulated by speaker goals.

Experiment 2

In Exp. 2, we examined whether a speaker's sensitivity to the perspective-status of contrast objects is influenced by utterance goals. Utterances can be used for various purposes (Searle, 1969), but in this experiment, we focused on two: informing and requesting. We hypothesized that speakers would be more likely to design expressions with respect to common ground when requesting, as the addressee must interpret the speaker's message precisely in order to complete the request. In contrast, interpretation of information in the absence of some explicit behavioral goal (e.g., to move something) may afford less precise, i.e., "good enough" understanding (e.g., Ferreira, et al., 2002). This hypothesis is motivated by a view of language use that multiple information sources constrain processing—crucially, including what the user is trying to accomplish with her language (Brown-Schmidt & Hanna, 2011).

Method

Participants Thirty pairs of students at Seoul National University who did not participate in Exp. 1 participated in return for cash payment. Pairs were all friends; one participated as speaker and the other as an addressee. While our analyses focus on the speaker's role, both participants were led to believe they played key roles in the study.

Materials The materials were similar to those used in Exp. 1. The main difference was that each participant received 11 objects per set, with 6 critical objects (3 targets and 3 competitors), plus 5 distractor objects. Subjects were presented with 18 critical and 18 filler trials. In addition, unlike in Exp. 1, the target was the larger of the two contrast objects ½ the time, and the smaller the other ½. Pairs of critical objects were placed near each other so participants could easily identify the two objects and appreciate that their perspective conflicted with their partner's.

Procedure The procedure was similar to Exp. 1. Two manipulations were used: Common ground (these three conditions were identical to Exp. 1), and utterance goal (informing vs. requesting). On informing trials, the experimenter was the actual agent moving the objects and the speaker provided information to the addressee about where the objects would be moved before the experimenter moved them (e.g., "*The experimenter will move the plate to the left*"). On requesting trials, the speaker asked the addressee to move objects (e.g., "*Can you move the plate to the left*?"), and the addressee, rather than the experimenter, moved the object. The common ground conditions were manipulated within subjects and utterance goal was manipulated between subjects. Participants were randomly assigned to the goal conditions; items were counterbalanced. Finally, unlike Exp. 1, participants did not switch roles with each other—speakers and listeners continuously played their respective roles throughout the entire experiment.

Coding The coding of scalar terms was identical to Exp. 1. Additionally, the speaker's spatial language was categorized by whether it was produced from the speaker's perspective (ex. "*right*" = speaker's right) or the addressee's perspective (ex. "*right*" = addressee's right).

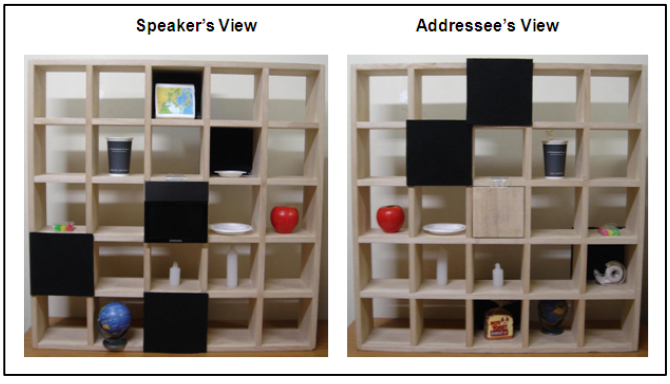

Figure 3. Example display. Critical objects include 1 cup, 2 candles and 2 plates; the small plate was hidden from the addressee.

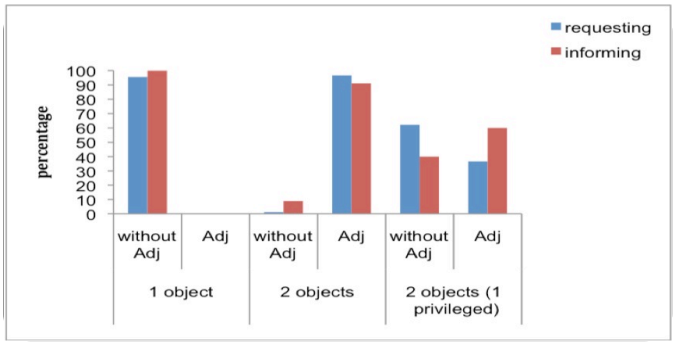

Figure 4. Target descriptions in Experiment 2.

Results and Discussion

In the common ground conditions (1-object and 2-objects), the results were similar to Exp. 1 ($p > .05$), regardless of utterance goal (informing vs. request), see Fig. 4. This result demonstrates that the goal manipulation did not simply change speakers' overall motivation in the task or attention to utterance form. In contrast, when speaker and addressee's perspectives differed (2-objects 1 Privileged), utterance goal modulated referring. Participants named the object without modification 62.22% when requesting vs. 40.00% when informing. Modification rates complemented this pattern, with 36.67% and 60.00%, modification rates in the requesting and informing conditions, respectively. A mixed-model analysis revealed a significant effect of goal on use of contrast adjectives (b=0.349, SE=0.0225, $p < .01$). When speakers made a request, common ground constrained language production more, resulting in fewer expressions designed with respect to privileged competitors. Thus, when speakers were in a situation in which they had to ask for something, speakers avoided adjectives that were unnecessary and potentially confusing from the addressee's perspective (see Wardlow Lane & Ferreira, 2008).

Furthermore, goals significantly affected the use of perspective in spatial language (Fig. 5). Recall that participants faced each other, thus the speaker's egocentric right was the addressee's egocentric left. We coded whether speakers used egocentric or addressee-centric spatial terms. When requesting, speakers always adopted the addressee's perspective, using expressions such as "*on your left*" when giving directions. In contrast, when informing, 5 of 15 participants consistently used egocentric spatial terms designed from their own perspective. The effect of goal on the use of egocentric spatial language was significant (b=0.349, SE=0.0225, $p < .01$). This result adds to the evidence that the process of tailoring an utterance for the addressee's needs depends on what the speaker is trying to accomplish with her utterance. When a speaker makes a request, the addressee's perspective has a stronger influence on utterance form than when the speaker informs.

Finally, one consideration is that perhaps these results are due to low engagement in the informing task. We suspect this is unlikely. In fact, error rates on informing and requesting trials show the opposite pattern, with 7 total errors on requesting trials and 2 total errors on informing trials. Thus, the higher dependence on common ground when requesting is likely due to differences in the relevance of perspective to goals when requesting, compared to informing.

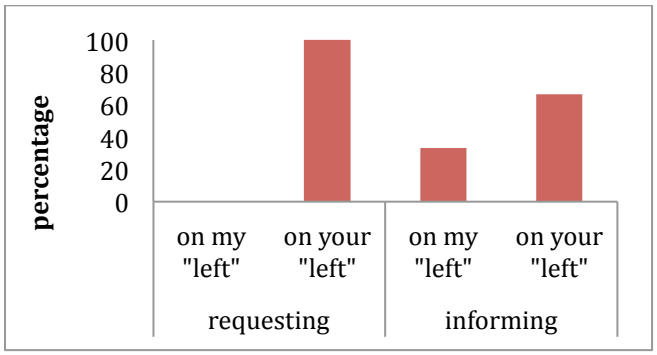

Figure 5. Spatial language in Experiment 2.

General Discussion and Conclusions

In Exp. 1, we examined whether speakers would use scalar modified referential expressions even when a scalar-contrast object was not visible to the addressee. When critical objects were in common ground, participants produced a bare noun when there was only one object, but produced a noun phrase with a scalar adjective when there were two mutually-visible different-sized objects of the same kind. In contrast, in the privileged condition, the contrasting object was blocked from the addressee's view, and speakers referred to the target with bare nouns on half of trials. The fact that adjectives were used significantly less often when the contrast was privileged demonstrates sensitivity to common ground in utterance design. However, common ground was not a complete constraint on production—after all, speakers used adjectives which were unnecessary from the addressee's perspective on half of the trials. These findings demonstrate—in a new paradigm, and a new population—that information from both the speaker's and addressee's perspective influences production.

Based on previous work with individuals from an interdependent culture (e.g., Wu & Keysar, 2007), we might have expected better use of common ground in this study. After all, Wu and Keysar found in a comprehension task, that Chinese, but not American participants, showed no significant interference from privileged lexical competitors. Why then, did our Korean participants, who are similar to Chinese in terms of an interdependent culture, fail to completely suppress privileged objects? One tantalizing possibility is that in language production (unlike comprehension), the speaker's privileged perspective is always relevant, as speaking often involves imparting new, previously privileged information to the addressee.

The results of Exp. 2 provide new key evidence that utterance goals influence a speaker's sensitivity to perspective. When critical objects were in common ground, goals played little role in referring, and results were similar

to Exp. 1. In contrast, when the contrast was privileged, the percentage of bare noun responses (the appropriate response from the addressee's perspective) was significantly higher when speakers were requesting vs. informing.

Why were speakers more sensitive to perspective when requesting? Here we suggest that when one's goal is to make a request, it is more critical that the addressee interpret the speaker's message precisely, in order for the requested action to be carried out correctly. In contrast, when informing, even if the addressee does not understand the speaker's meaning precisely, this would not likely have direct influence on the speaker. In some sense, then, the heightened sensitivity to addressee perspective when the goal is to request is a form of egocentrism—when the speaker wants the addressee to do something for her, she takes extra care to make sure that her request is understood.

The findings presented here provide tantalizing initial evidence for this view. When a speaker makes a request, the addressee's perspective has a stronger influence on utterance form than when the speaker informs the addressee. Thus, the relevance of the addressee's perspective to utterance goals circumscribes the degree of audience design. These results allow two important conclusions: First, the addressee's ability to uniquely identify the intended referent is not necessary for communicative success. Instead, communicative success must be interpreted with respect to the speaker's goals. Here we have shown that the addressee's ability to uniquely identify is less relevant when the listener is being informed. Second, investigations of the *ability* of speakers to engage in audience design must consider whether audience design is, in fact, relevant to communicative goals. Findings of audience design failures in situations in which audience design is not necessary to communicative goals are likely to be uninformative about the ability of speakers to engage in audience design.

Current models of the role of perspective in language production propose that the addressee's needs are only relevant at a delayed stage of processing (Horton & Keysar, 1996), or only operate as a weak constraint (Wardlow Lane et al. 2008). The results of the present research paint a much more complex picture. Common ground is clearly not an absolute constraint on production (even for individuals from an interdependent culture), however it does have clear and consistent effects. More importantly, we have demonstrated that understanding how and when common ground guides referring will crucially require an appreciation of what the speaker is trying to do with her utterances (i.e., Searle, 1969). Here we have taken a first step in this direction, demonstrating that the relevance of the addressee's perspective to utterance form is circumscribed by utterance goals.

References

Baayen, R.H., Davidson, D.J., & Bates, D.M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. Journal of Memory and Language, 59, 390-412.

Brown-Schmidt, S., & Hanna, J. E. (2011). Talking in another person's shoes: Incremental perspective-taking in language processing. Dialog and Discourse, 2, 11-33.

Epley, N., Keysar, B., VanBoven, L.V., & Gilovich, T. (2004). Prespective Taking as Egocentric Anchoring and Adjustment. Journal of Personality and Social Psychology, 87, 327-339.

Epley, N., Morewedge, C. K., & Keysar, B. (2004). Perspective taking in children and adults: Equivalent egocentrism but differential correction. *Journal of Experimental Social Psychology,* 40:760-768.

Ferreira, F., Bailey, K. G. D., & Ferraro, V. (2002). Good-enough representations in language comprehension. Current Directions in Psychological Science, 11, 11–15.

Grice, H. P. (1975). Logic and conversation. In Core P., & Morgan, J.L. (Ed.), Syntax and semantics 3: Speech arts (pp.41-58). New York: Academic Press.

Hanna, J. E., Tanenhaus, M. K., & Trueswell, J.C. (2003). The effects of common ground and perspective on domains of referential interpretation. Journal of Memory and Language, 49, 43-61.

Heller, D., Gorman, K., & Tanenhaus, M. K. (in press). To name or describe: shared knowledge affects referential form. *TopiCS.*

Horton, W.S., & Keysar, B. (1996). When do speakers take into account common ground? Cognition, 59, 91-117.

Keysar, B., Barr, D. J. Balin, J. A., & Paek, T. S. (1998). Definite reference and mutual knowledge: Process models of common ground in comprehension. *Journal of Memory and Language,* 39, 1-20.

Keysar, B., Lin, S., & Barr, D. J. (2003). Limits on theory of mind use in adults. Cognition, 89, 25-41.

Nadig, A. S., & Sedivy, J. C. (2002). Evidence of perspective-taking constraints in children's on-line reference resolution. Psychological Science ,13, 329- 336.

Olson, D. R. (1970). Language and thought: Aspects of a cognitive theory of semantics. Psychological review, 77, 257-273.

Searle, J. (1969). Speech Acts. Cambridge University Press.

Tanenhaus, M., & Trueswell, J. (1995). Sentence comprehension. In J. Miller & P. Eimas (Eds.), Handbook of cognition and perception: Iiol. 11. Speech and Language. San Diego, CA: Academic Press.

Wardlow Lane, L., Groisman, M., Ferreira, V. S. (2006). Don't talk about pink elephants! Speakers' control over leaking private information during language production. Psychological Scienc, 17, 273–277.

Wardlow Lane, L., & Ferreira, V. S. (2008). Speaker-External versus speaker-internal forces on utterance form: Do cognitive demands override threats to referential success?

Wu, S., & Keysar, B. (2007). The effect of communication overlap on communication effectiveness. Cognitive Science, 31, 169-181.