Compensating for an Inattentive Audience

By

Nicole Neekoo Craycraft

Thesis

Submitted to the Faculty of the

Graduate School of Vanderbilt University

in partial fulfillment of the requirements

for the degree of

MASTER OF SCIENCE

in

Psychology

May, 2017

Nashville, Tennessee

Approved: Dr. Sarah

Brown-Schmidt, Ph.D.

Dr. Megan Saylor, Ph.D.

ACKNOWLEDGMENTS

This material is based on work supported by National Science Foundation Grants NSF BCS 12-57029 and NSF BCS 15-56700 to Sarah Brown-Schmidt. Thank you to Sarah Brown-Schmidt for the invaluable guidance and support throughout grad school. Thank you to Mahima Goel for help with transcriptions and participating as a confederate, and to all other undergraduate research assistants who have helped with this work.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

Over the course of a lifetime, we have conversations in a variety of different contexts and circumstances, and often successfully communicate despite potential challenges these situations might impose on conversation. Here we focus on the challenge of communicating with an inattentive conversational partner. In a situation where information must be communicated to a partner, how do speakers achieve successful communication when that partner is inattentive?

According to a prominent view, the common ground held between conversational partners – the knowledge and beliefs that they mutually hold – forms the basic context for language use (Clark, 1996). Representations of what is and what is not common ground are critical to basic processes such as asking an informational question (i.e., asking about something not in common ground; Brown-Schmidt, Gunlogson, & Tanenhaus, 2008), or describing one of several mutually visible items to a co-present listener (Clark & Wilkes-Gibbs, 1986). A classic finding is that when conversational partners repeatedly discuss entities that lack an established name (e.g., an abstract image), over the course of the conversation they form brief, efficient labels for these items, reflecting the formation of common ground (Krauss & Weinheimer, 1964; 1966; Clark & Wilkes-Gibbs, 1986). Critically, these brief labels reflect common ground held with that *particular* conversational partner. In studies where a speaker develops brief labels for a series of

abstract images with one partner, and then switches to a new partner who has no experience with the labels, speakers tend to revert to long descriptions (Wilkes-Gibbs, & Clark, 1992; Holler & Wilkin, 2009; Horton & Spieler, 2007; Horton & Gerrig, 2002; 2005; also see Brennan & Clark, 1996; Galati & Brennan; 2010). Similarly, in conversations among three parties, where a speaker must alternate between addressing one partner who is knowledgeable about image labels, and a second partner who is naïve, speakers successfully produce brief labels for the knowledgeable partner, and longer descriptions for the naïve partner (Yoon & Brown-Schmidt, 2014; submitted). These longer descriptions are necessary in order for the naïve partner to successfully understand the speaker; indeed, overhearers have a difficult time understanding these brief labels once they have been established (Schober & Clark, 1989; Kraut, Lewis, & Swezey, 1982; also see Fussell & Krauss, 1989).

This process of tailoring what is said based on what a specific listener does and does not know is known as *audience design* (Clark & Murphy, 1982; Fussell & Krauss, 1989), and is thought to require access to representations of common ground that are bound in memory to specific conversational partners (Brown-Schmidt, Yoon, & Ryskin, 2015; Horton & Gerrig, 2005). For audience design to be successful, speakers must make accurate inferences about what their communication partner does and does not know. What sources of information then, do speakers use to infer what addressees do and do not know? Clark and Marshall (1978; 1981) propose that conversational partners use simple heuristics to make inferences about what is and is not common ground based on evidence in the communicative context. For example, imagine a speaker were to label a picture of a drooling monster for an addressee, "*This one is called Banpar"*. According to Clark and

Marshall's theory, the speaker could assume that the monster, and the label that goes with it, are common ground if the addressee is present, attending to what is said, and rational (the *simultaneity, attention,* and *rationality* assumptions, respectively).

Findings that speakers do not assume common ground with listeners who are not actively involved in a conversation -- such as bystanders or side participants (Wilkes-Gibbs & Clark, 1992) support the hypothesis that simultaneously experiencing an event contributes to inferences about common ground. Similarly, when the speaker and addressee cannot freely interact, and the speaker lacks full access to the addressee's feedback, communication suffers (Clark & Krych, 2004), indicating that feedback may play an important role in formation of common ground (Clark & Bangerter, 2004; Clark & Wilkes-Gibbs, 1986; Roque & Traum, 2008; cf. Brown-Schmidt, 2012). Whether the *attentiveness* of a co-present conversational partner guides inferences about common ground is less clear. Matarazzo, et al. (1964) find that in a clinical employment interview setting, civil service applicants use longer conversational turns when the interviewer nods throughout the conversation. Studies of narrative production also suggest that under some circumstances, speakers address attentive and inattentive listeners differently. For example, Kuhlen and Brennan (2010) asked speakers to tell jokes to a listener, and measured the amount of extra detail they added to their delivery of the joke. Speakers were either told to expect an attentive addressee, or a distracted addressee; whether the addressee was in fact attentive or distracted was independently manipulated. Kuhlen and Brennan found that jokes were told with about the same amount of detail across these situations except for the case where speakers expected, and experienced, an attentive addressee – in this case, they added more detail to their jokes. Similarly, Pasupathi, Stallworth, and Murdoch (1998) find

that speakers elaborate spoken narratives more for attentive than inattentive addressees (also see Bavelas, Coates, & Johnson, 2000; Kraut, et al. 1982). These findings suggest that speakers can detect inattention on the part of an addressee, and that inattention can, in turn, affect some aspects of language production.

What these findings do not address, however, is whether the attentiveness of a partner modulates assumptions about common ground as hypothesized by Clark and Marshall (1978; 1981). While we know of no evidence that speaks directly to this hypothesis, related work examining common ground-based processes in language comprehension points to a rather limited role for addressee feedback. Using measures of on-line language comprehension, Brown-Schmidt (2009) found that when a partner explicitly denied hearing what was said, e.g., "*Sorry I didn't hear you"* that what had previously been said was not assumed to be common ground. However, subsequent research manipulating different types of listener feedback hypothesized to provide different amounts of evidence for common ground (e.g., repeating back what was said, vs. "OK", vs. silence) found little to no effects (Brown-Schmidt, 2012; Brown-Schmidt & Fraundorf, 2015; cf. Clark & Schaefer, 1989). Taken together, these findings suggest that so long as the addressee does not flat-out deny hearing what was said, that simply being co-present in a situation may be sufficient to license an assumption that what was said is common ground. On the other hand, findings that narrative production is influenced by addressee attentiveness (Kuhlen & Brennan, 2010; Pasupathi, et al., 1998; Bavelas, et al., 2000) suggest that speakers are sensitive to the addressee's behavior in the moment. Whether this influences inferences about addressee knowledge (and thus what is assumed to be common ground between the conversational partners), remains an open question. Another

unanswered question concerns the specificity of this influence: if an addressee exhibits signs of distraction at one point in the conversation, do speakers assume that they missed what was said during the period of distraction alone, or is the effect of distraction more general, leading speakers to infer that the addressee may have missed other information in the conversation as well?

Here we test the hypothesis that in conversational situations, speakers only assume common ground for what they have said when the co-present addressee exhibits *attentiveness* in the conversational situation. A relevant real-world example is as follows: Imagine I am talking to a friend about my favorite character in the new TV show I'm watching. If she is listening intently, I can feel secure in the knowledge that she knows the character's name, however, if she pulls out her phone as I am speaking, I am left unsure of how much she actually heard, and thus I may be less likely to assume her familiarity with the character's name in later conversation. Across three experiments, the present research explores the role of partner attentiveness in formation of conversational common ground, and the specificity of this relationship. We examine situations in which conversational partners learn the names of a series of cartoon monster pictures together – a paradigm that is known to elicit formation of common ground for the names (Gorman, et al., 2013; also see Heller, et al., 2012). Manipulations of addressee attentiveness before or during the name-learning task are used to evaluate the role of attentiveness in formation of common ground, and the specificity of this relationship. Based on well-established findings that speakers use longer descriptions of images when addressing a naïve partner who lacks common ground for the image names (Wilkes-Gibbs & Clark, 1992; Isaacs & Clark, 1987; Yoon & Brown-Schmidt, 2014; Horton & Gerrig, 2005), we use referential description

length as a tool to evaluate whether speakers take into account addressee attentiveness when gauging whether common ground has been formed, or not.

CHAPTER 2

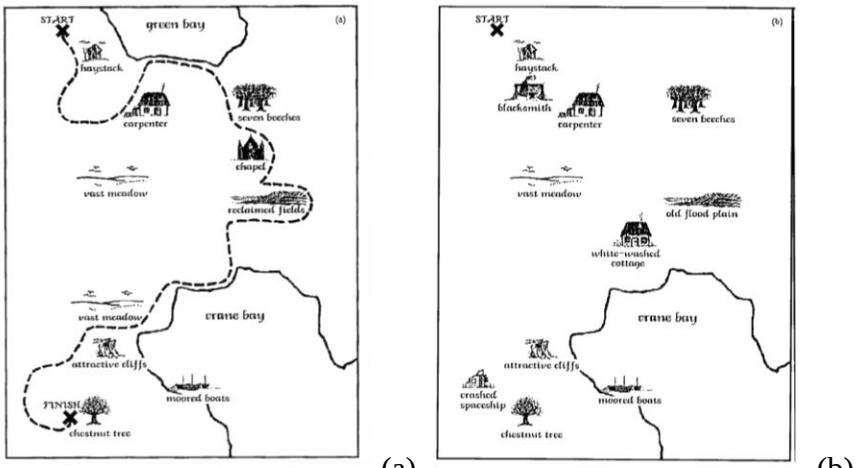
EXPERIMENT 1

In Experiment 1 we examine situations in which speakers interact with two conversational partners, one of whom displays signs of inattention early in the conversation. We ask whether this lack of engagement has non-specific (general) effects on assumptions about common ground for information that is presented later in the conversation.

2.1 Methods

Participants

30 pairs (N = 60) of native English speaking undergraduates from the University of Illinois Urbana-Champaign served as participants in exchange for partial course credit. An additional 7 pairs (N = 14) were run and excluded from analysis for scoring below the accuracy threshold (5) and experimenter error (2); see details below.



Figures 1a-b. Example stimuli from the map task for the speaker (1a) and listener (1b), from Anderson et al. (1991). Note that, by design, the maps contain discrepancies.

Procedure

Participants completed the experiment in pairs. At the beginning of the study, the pair of participants were introduced to each other and were told that a lab assistant would also be participating. The participating lab assistant was introduced as never having done the task before and currently participating so as to be trained on the task in order to run other participants in the future. This cover story was never questioned by any of the true participants. An overseeing experimenter observed all trials of all the tasks. One of the true participants was randomly selected to be the speaker, and the other true participant and the participating lab assistant played the role of matchers.

There were three phases to the task. In the first phase of the task, the speaker and the two matchers were seated in a room together and each was given a paper copy of a map. They were seated so that they could not see each other's maps. The speaker and the two matchers completed a version of the Map Task (Brown, Anderson, Yule, and

Shillcock, 1983). The speaker's map had a path drawn between landmarks (Figure 1a), whereas the two matchers each had a copy of a similar map that did not contain the path (Figure 1b). The presence and names of landmarks on the maps were not identical and participants were not informed of this ahead of time, which increases the difficulty of the task. The speaker was instructed to describe the path to the matchers so that they could draw the same path on their map. Matchers were instructed to pay attention and to ask questions if needed. The matchers could see each other's maps but were not explicitly informed their maps were identical.

The lab assistant matcher behavior during the map task was manipulated in a between-subjects manner. In the Engaged condition, the lab assistant matcher participated fully and was engaged. In the Disengaged condition, the lab assistant matcher was completely inattentive during this task, actively not participating at points, and generally disengaged from participation. To standardize the behavior in each condition, a checklist of behaviors to complete was memorized by all lab assistant matchers. Disengaged behaviors include checking a cell phone, not making eye contact, looking around the room, and making a mistake on the map path. Table 1 presents differences in this scripted behavior in the two conditions. The overseeing experimenter had a clipboard and took copious (unnecessary) notes, to provide a plausible explanation for why they did not notice this behavior in the Disengaged condition. The task was completed when the participants notified the overseeing experimenter that the paths were completed, who then confirmed if the matchers' paths matched the speaker's.

Fifteen pairs of participants were randomly assigned to the Engaged condition and fifteen pairs were randomly assigned to the Disengaged condition.

Table 1
Participating lab assistant behaviors, by condition

Engaged	*Disengaged*
Physical behaviors	
Sit up straight in chair	Slouch
Maintain eye contact with whoever is speaking	Don't make eye contact with anybody
Look at the current talker	Look at the table, instead of the talker
Hands engaged in task, or quietly on table/lap	Playing with hair/ picking at nail polish etc.
Verbal behaviors	
Asking questions after speaker finishes talking	Asks questions after multiple conversation turns
Promptly confirms understanding when asked	Is silent initially, answer is delayed when asked
Social behaviors	
Doesn't use distractions during task	Distracted during task (e.g., pulling phone out)
Shows engagement in task only	Glancing around room multiple times
Neutral, small smile facial expression	Bored, flat facial expression

In the second phase of the task, the participants were seated at three separate computers in the same room that they had completed the Map Task in. The task in this phase was to learn the names of a series of 12 unfamiliar monster images. The monsters, which were presented on the computer displays, were a subset of the stimuli used in Gorman et al. (2013).

The participant roles (speaker, matcher, lab assistant matcher) from the Map Task were maintained in the name learning task. All participants were instructed to learn the names of the novel monsters and were told that they would be tested on the names later. A round of name learning consisted of two stages: group name learning, and then individual name review.

The speaker was seated at the end of one table, with the matchers on the other side

of the table. In the group name learning stage, the speaker rotated their laptop so that all three members of the group could see the screen. The setup of the group name learning stage is pictured in Figure 2.

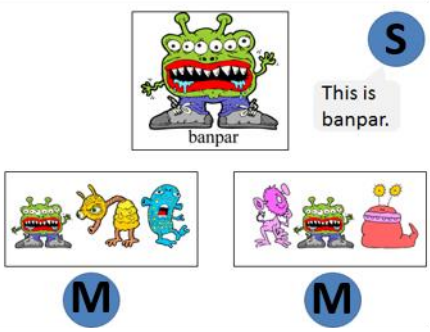


Figure 2. The setup for the group name learning (Where S is the speaker and M is a matcher). All participants could see the speaker's screen, but the speaker could not see the screens of the matchers.

The matchers were able to see both their own individual laptops, as well as the director's laptop screen. On each trial, the three participants viewed an image of a monster with the name written underneath it on the director's laptop. The speaker was instructed to read out loud the name of that monster, to ensure the names were attended to and to increase the likelihood of the speaker remembering that monster's name (cf. the generation effect; Slamecka & Graf, 1978, Bertsch, Pesta, Wiscott, & McDaniel, 2007). On the matchers' laptops, three unlabeled monsters were displayed: the target monster that matched the one on the speaker's screen and two randomly selected distractor monsters. The matchers were instructed to click on the target monster on their individual laptops. After both matchers clicked, the speaker clicked their own laptop to move to the next trial. The 12 monsters were shown, one at a time, in a random order that was the same for all

groups, with each monster shown twice in a round. Thus the participants had two chances to learn each monster's name.

Unlike the first phase of the task, the behavior of the lab assistant matcher was held constant across conditions during group name learning. In both the Engaged and Disengaged conditions, the lab assistant matcher was instructed to participate without comment and not engage in overtly inattentive behaviors. Since the speaker could not see the matchers' screens, it was left ambiguous as to whether the lab assistant matcher was truly selecting the correct monster, or just clicking at random.

Following group name learning was a self-paced review of the names. All participants completed this individual self-paced review. Participants stayed in the same room, and the rotated laptop was turned back to face the speaker. During this review, the 12 monsters were shown on the screen, one at a time, unlabeled, and in a different random order than during the group name learning, which was the same across all participants and groups. The participants were instructed to mentally rehearse the monster's name, and then click the mouse. The monster then appeared labeled on the screen so that the participant was able to check if they correctly remembered the name. The silent and individual nature of this task was intended to keep ambiguous whether the lab assistant was truly learning the monster names, while also providing an opportunity for the true participants to assess which names they had learned effectively. The individual review phase is pictured in Figure 3. A round of the name-learning task consisted of the name learning stage (repeated twice) and one repetition of the self-paced review. The name-learning task lasted for eight repetitions of the group name learning and the self-paced review, thus all together participants were given 24 exposures to each monster name with 8 additional self-test

opportunities.

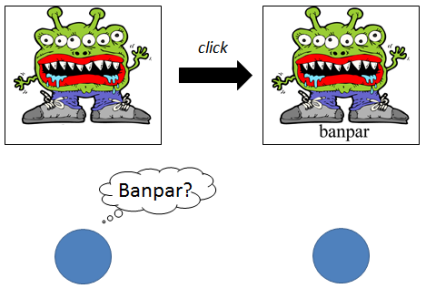


Figure 3. The setup of the individual name review. Participants were self-paced, with no time limit per name.

The third phase of the task was a referential communication task (Krauss & Weinheimer, 1966) in which the speaker gave the matchers (one at a time) instructions to re-arrange a series of pictures. It was during this phase of the task that we collected the data of primary interest for analysis. The speaker interacted with each matcher individually while the other matcher was in another room. The order in which the speaker interacted with the two matchers was counterbalanced across subjects. The speaker and the matcher used separate laptops and remained in the same room as the previous tasks. The speaker's display had a set arrangement of 16 monsters (12 named, 4 novel) and the matcher's display had the same monsters but in a randomized order. The novel monsters were included to increase the difficulty of the task. An example speaker's display is seen in Figure 4. The speaker's task was to give the matcher instructions for how to re-arrange the monsters so that the arrangement of the monsters on their screens matched. The matcher followed the speaker's instructions by dragging the monsters around the screen with the mouse. The speaker was not given explicit instructions for how to refer to the monsters; if asked directly, the overseeing experimenter said that they were welcome to say whatever

they wanted. Matchers were told that they could ask questions if needed and could freely communicate with the speaker. Matchers mostly provided positive feedback when successfully positioning an item (e.g. *mm-hm*, *okay*), and rarely asked questions. When matchers did ask questions, it was typically to clarify the position of a monster (e.g., *Wait that one was under bluchek?*) or to seek confirmation (e.g., *Pink one right?*). The lab assistant matcher was instructed to fully engage in this task, regardless of condition. The task was completed when all the monsters were arranged correctly and checked by the overseeing experimenter.

Once the speaker and the first matcher completed the referential communication task, the two matchers switched places and the speaker repeated the task with the second matcher. In this second round, the order of the pictures was changed and the novel monsters were swapped with a set of new novel monsters. It was during this phase in which the conversation was recorded for analysis. The audio was recorded using the microphone on the speaker's Macbook laptop. Both the speaker and the matcher's utterances were transcribed.

Once both matchers participated in the task, all participants were given an individual post-test. Participants were handed a piece of paper with the twelve critical monsters pictured and unlabeled, and they had to correctly write down the names of all the monsters. Overall, the entire experiment took about forty minutes.

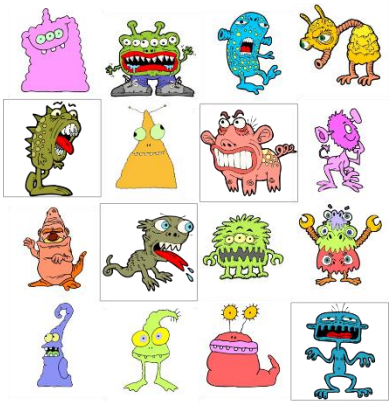


Figure 4. Example of a speaker's display during the referential communication task. Novel monsters are boxed in gray. The corresponding matcher's screen had the same 16 monsters randomly positioned.

Predictions

Success in the critical experimental task – the referential communication task – requires the speaker to provide enough details for the addressee to identify each intended referent; doing so requires the speaker to make assumptions about the addressee's knowledge (Wilkes-Gibbs & Clark, 1992). This allows us to make specific predictions about the relationship between assumptions about common ground and referential form in this task. Based on previous studies of audience design with these stimuli (Gorman, et al., 2013), we expect speakers to produce longer descriptions of the monsters if common ground for the names cannot be assumed. If addressee engagement has nonspecific effects on subsequent inferences about common ground, speakers should be less likely to assume common ground for the names of the monsters, and produce longer descriptions (include more information) for addressees who were previously disengaged. Such a result would obtain if speakers assumed that previously disengaged addressees were unlikely to have learned the names of the monsters, and required more descriptive information in order to

identify them. Alternatively, if addressee feedback has limited effects on assumptions about common ground (Brown-Schmidt, 2012; Brown-Schmidt & Fraundorf, 2015), or these effects are restricted to the specific information exchanged during the period of distraction, speakers should describe the monsters in a similar way for the disengaged and engaged matcher.

2.2 Results

Data Exclusions

Only the groups in which the speaker scored at or above 80% accuracy on the post-test of the monster names were included in the analysis (5 groups of participants were excluded based on this criterion). This accuracy criterion was used to ensure that the speaker had the capability of naming the monster when appropriate. Mean accuracy on the post-test for the final group of 30 pairs of participants included in the analysis was 95.8%.

Our analyses focus on utterances produced by the speaker, as our hypotheses concern how the attentiveness of a conversational partner influences the process of audience design. In particular, we focus our analyses on the length of the descriptions that speakers used to describe the monsters that they had learned the names for in the name-learning phase. Novel referents were excluded from analysis because no differences between conditions were expected for them. Due to the nature of the picture-arranging task, it is possible for the matcher to place the final monster using a process of elimination (see Clark &Wilkes-Gibbs, 1986). For example, sometimes the matcher will provide the name (*and the last one is inta?*) and the speaker just affirms (*Yup good*). Because of this,

the last referent discussed was excluded from analysis. Any trials where the speaker used an incorrect name were excluded, eliminating 1.6% of data.

Analysis

The primary dependent measure was the total number of words used by the speaker to describe each monster. For example, a description of *danzo the blue monster with the yellow polka dots* would have a length of 9. If the speaker just used the name (e.g., *The eighth character is nelky*), this was counted as a description length of 1. Lexical disfluencies (e.g., "um") were not included in the description length measure.

To examine the effect of partner engagement on the length of speakers' utterances, we used a mixed effect linear model, using the Poisson distribution due to the count nature of the data. The model was fit with the maximal random effects structure for subjects and items. When the maximal model did not converge, a backwards-fitting procedure was implemented to find the model with the largest random effects structure that successfully converged (Barr et al., 2013). Recall that in both the Engaged and the Disengaged conditions, speakers addressed two different Matchers. For speakers in the Engaged condition, both Matchers were engaged (the participant Matcher and the lab assistant); for speakers in the Disengaged condition, one Matcher was engaged (the participant Matcher) and one was disengaged (the lab assistant). Note that while we consider participant Matchers as "engaged" for the purposes of analysis, we did not explicitly manipulate their behavior (though they were generally engaged). The predictors used in the model were factors for Matcher Type (lab assistant matcher or true matcher) and Condition (Engaged or Disengaged) and Trial (first or second time the speaker completed the referential

communication task). This factor was included based on previous findings that repeated mention to a given item tends to result in shorter utterances (Wilkes-Gibbs & Clark, 1992; Yoon & Brown-Schmidt, submitted). All predictors were mean-centered.

Table 2 shows the results of the model. There was a significant main effect of Trial, such that speakers used shorter descriptions when describing the monsters with the second matcher (β = -0.40, $p < .05$). There was also a significant main effect of Condition: speakers used longer descriptions in the Disengaged condition compared to the Engaged condition, independent of who the speaker was currently addressing (β = 0.29, $p < .05$). There was no main effect of Matcher Type (participant or lab assistant) indicating that speakers produced similar length utterances for both types of addressees (β = 0.05, $p > .6$). The predicted interaction between Matcher Type and Condition was not significant ($p > .7$). Figure 5 shows the average description length for each matcher type in each condition.

Table 2

Full mixed effects model for Experiment 1

	Estimate	SE	z-value			Var.	Stand. Dev.
Fixed				*Random*			
Trial	-0.397	0.183	-2.167*	*Subject*	(Intercept)	0.116	0.341
Matcher Type	0.045	0.112	0.406		Trial	0.299	0.547
Condition	0.298	0.135	2.207*	*Item*	(Intercept)	0.182	0.426
Trial : Matcher Type	-0.081	0.270	-0.301		Trial	0.223	0.473
Trial : Condition	-0.107	0.225	-0.478				
Matcher Type : Condition	-0.059	0.225	-0.262				
Trial : Matcher Type : Condition	0.730	0.539	1.354				

Note: * = significant at p < .05, ** = significant at p < .01, *** = significant at p < .001.

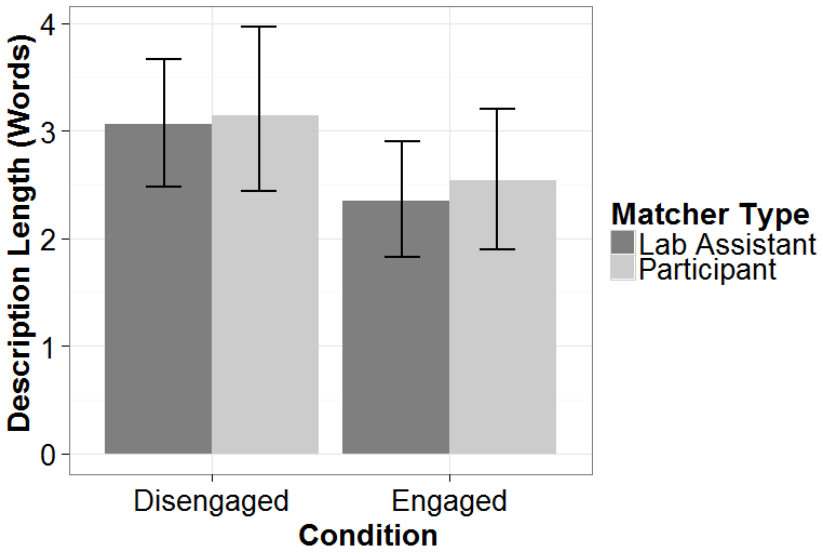


Figure 5. The average description length (in words) for the different types of matchers in the different conditions. In the Disengaged condition, the lab assistant matcher was disengaged from the task; in the Engaged condition, the lab assistant matcher participated in the task fully. The behavior of the participant Matcher was not experimentally manipulated.

2.3 Discussion

We hypothesized that the engagement of a conversational partner would influence the speaker's assumptions about what information was in common ground with that partner. We predicted that if listener attentiveness has general effects on assumptions about common ground, speakers would label objects using more words when addressing the partner who had been previously disengaged during the map task. Speakers were in fact more descriptive overall in the Disengaged condition: however, they did not differentiate between the previously engaged and disengaged partners. Thus, there was no evidence that

speakers took into account an addressee's previous level of engagement (during the map task) when making assumptions about that specific addressee's knowledge in a subsequent task (during the name learning task).

One explanation of these findings is that speakers assume what they say to another person is common ground with that person, regardless of the type of feedback that listener provides (Brown-Schmidt, 2012; Brown-Schmidt & Fraundorf, 2015). An alternative possibility – one we will explore in Experiment 2 – is that the (in)attentiveness of an addressee *does* influence inferences about what is and is not common ground, but that these inferences are specific to the information that was exchanged during period of inattention. By design, in the Disengaged condition, the disengaged matcher did not produce any overt behaviors to positively indicate a lack of engagement in the period of time following the map task. Speakers, then, may have assumed that this prior inattentiveness did not compromise the disengaged matcher's ability to learn the monster names in the subsequent conversation. In Experiment 2 we test this hypothesis by manipulating addressee engagement during the name-learning task itself to test for specific effects of addressee engagement on inferences about what is and is not common ground.

A final question concerns why speakers produced longer utterances overall when the lab assistant matcher had been disengaged. One explanation is that the experience of interacting with a disengaged matcher increased overall uncertainty in the situation. To be sure, however, the fact that speakers used similar expressions for disengaged and engaged matchers at test, is inconsistent with the hypothesis that inattention affected assumptions about common ground.

CHAPTER 3

EXPERIMENT 2

Experiment 2 was similar to Experiment 1, but with some major modifications. First, because there were no differences in how speakers addressed the true participant and lab assistant participant, in Experiment 2 only lab assistants served as matchers. Second, the overseeing experimenter left the room during the tasks in order to increase the plausibility of the disengaged behavior not being acknowledged. Third, we dropped the Map Task and instead manipulated the lab assistant's behavior during the name learning task. The critical referential communication task immediately followed the name-learning task, and we altered the referential communication task in order to allow both matchers to participate at the same time. Experiment 2 is designed to test for specific influences of addressee attention on assumptions about common ground; to do so we manipulated the lab assistant's behavior during a task – the name learning task – that was relevant to the critical referential communication task. If speakers use a conversational partner's (in)attention to assess common ground for what was said during that period of (in)attention, and this common ground (or lack thereof) guides subsequent audience design, then speakers should produce longer descriptions when addressing a partner who was previously disengaged, compared to a partner that was previously engaged.

3.1 Methods

Participants

33 native English speakers from the University of Illinois community served as participants for monetary compensation at a rate of $8 an hour. All participants were young adults aged 18 or older, and none had participated in Experiment 1. An additional 11 participants were run and excluded from analysis for scoring below the accuracy threshold.

Procedure

The participants completed the task in a group of three, however, unlike Experiment 1, the two other individuals in the group were lab assistants. The same cover story (that the participating lab assistants had never done the task before and are currently participating so as to be trained on the task in order to run other participants in the future) from Experiment 1 was used to explain the participation of lab assistants in the study. Experiment 2 consisted of two tasks: a group name-learning task and a referential communication task. The same monster stimuli from Experiment 1 were used in Experiment 2.

During the group-name learning task, the three partners in the group were told to work together to learn the names of the different monsters, and that they would be tested on the names later. The true participant was always "randomly" assigned to the role of speaker, and the two lab assistants were assigned to be the matchers. After giving the three partners the task instructions, the overseeing experimenter left the room, citing concerns over space. One laptop was positioned such that all partners could see the screen and the

participant speaker could control the pace of the task. A schematic of the room layout is seen in Figure 6.

During the group-name-learning task, the group first saw instructions were displayed on the laptop, saying that the group needed to learn the names of the different monsters together, and to click the mouse to move between trials. Then, there were 24 trials in which the 12 monsters were presented one-at-a-time, labeled with their names; each monster was seen twice and trials were in a random order. Next, participants saw instructions to work together to generate the name, and to click to move on after everybody said the name out loud. After these instructions appeared on the laptop, then the monsters were presented unlabeled, one-at-a-time, in a random order six times each, resulting in a total of 72 of these name-review trials. They had to work together to produce the name; upon clicking, the monster appeared with its label, allowing the participants to assess the accuracy of their initial guess. All participants were encouraged to say the names out loud, and the two lab assistants had been previously instructed to act as if they had never seen the monsters before. It was during this part of the task that we manipulated one of the lab assistant's behavior. In the Engaged condition, both lab assistants were fully attentive and participating, and in the Disengaged condition one assistant was engaged and the other was disengaged, implementing the relevant behaviors as described in Experiment 1. Altogether, participants were exposed to each monster-name pairing eight times before the test phase.

The final task was a referential communication task. Participants remained in the same room and in the same positions. As in Experiment 1, this task was the focus of our primary analyses and was audio-recorded. Each of the three partners was given a laptop by the overseeing experimenter, who then left the room again. The two matchers could see

each other's screens, but the speaker could not see the matchers' screens. On each trial, the three partners saw the same 4 monsters on their respective computers, and the participant speaker saw a box around one of the 4 monsters to indicate the target. A cue at the top of the speaker's screen indicated who the addressee was for that trial, either "Matcher A" or "Matcher B"; which matcher was the addressee randomly alternated on each trial. The speaker's task was to tell the indicated Matcher to click on the target picture; the other Matcher simply clicked to move forward to the next trial (see Yoon & Brown-Schmidt, submitted, for a similar approach). Example displays are shown in Figures 7a-b. The individual laptops used by matchers had large "Matcher A" and "Matcher B" labels that were visible to the speaker. Speakers were allowed to describe the target in any way, with the exception of location information (e.g., *top left*), as the position of the monsters was randomized for each participant separately. Both Matchers, regardless of their previous engagement level, participated fully. During this phase of the task, the speaker identified each monster twice for each matcher (to increase the amount of data), resulting in 48 trials. There were no filler trials. Unlike Experiment 1, there were no new monsters visible at test; this was done to prevent descriptions of the named monsters by comparing them to the novel monsters.

The same post-test of generating the names for the monsters from Experiment 1 was given to all participants after completion of the matching task for the same purpose of confirming participants had the capacity to use the names.

Participants were randomly assigned to the Engaged condition (n=17) or the Disengaged condition (n=16), each participant was assigned to one list. The pairings of names and monster illustrations was counterbalanced across 4 different lists. Overall, the

entire experiment took 30 minutes.

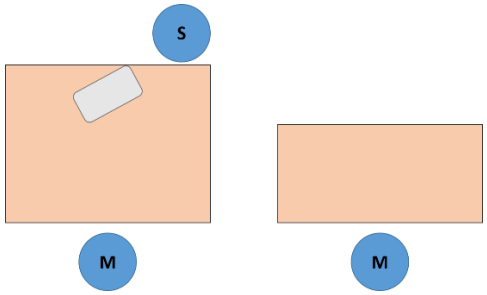


Figure 6. A schematic of the room for the name learning task.

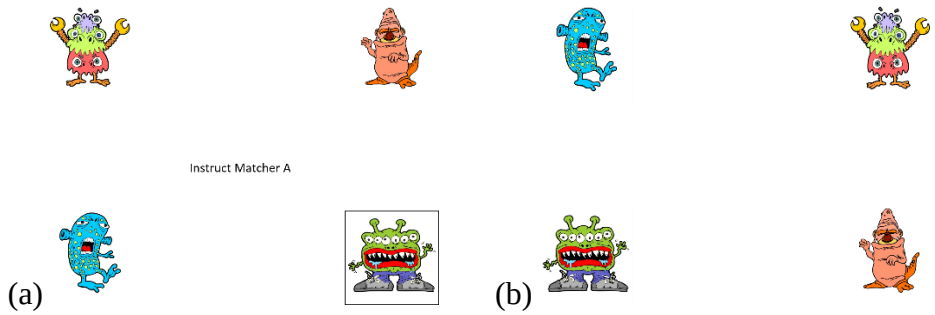


Figure 7. (a) Example speaker's display for the matching task. (b) Example matcher's display for the matching task.

Predictions

The manipulation of addressee engagement in Experiment 2 was during a phase of the conversation directly relevant to the critical referential communication task – whereas engaged matchers exhibited behavior consistent with learning the names of the monsters, disengaged matchers behaved as if they did not. If assessment of addressee knowledge for

the purposes of audience design is sensitive to addressee engagement, speakers should produce longer descriptions for addressees who were previously disengaged. Alternatively, if the behavior of the addressee has only limited effects on inferences about common ground (Brown-Schmidt, 2012; Brown-Schmidt & Fraundorf, 2015), speakers should produce similar descriptions for engaged and disengaged matchers.

3.2 Results

Data Exclusions

The same accuracy threshold from Experiment 1 of 80% performance on the post-test for all speakers was used, resulting in exclusion of 11 participants. The accuracy of the remaining 33 participants was 93.2%. As in Experiment 1, only the speaker's utterances were analyzed and trials where the speaker used the incorrect name at test were excluded, removing 0.1% of trials.

Analysis

The total word count was calculated in the same way as in Experiment 1. Like Experiment 1, the word count data were analyzed in a mixed effect linear model, using the Poisson distribution. Recall that in both the Engaged and the Disengaged conditions, speakers addressed two different lab assistant Matchers; in the Engaged condition, both Matchers were engaged, whereas in the Disengaged condition, one Matcher was engaged and one was disengaged. To distinguish between the two lab assistant Matchers in our analyses, in the Disengaged condition, the disengaged Matcher was coded as the critical matcher and the engaged Matcher as the control matcher. In the Engaged condition, both the critical and control Matcher were engaged; for the purposes of analysis, the two

Matchers were randomly assigned to the roles of critical and control. Thus, the predictor variables were Condition (Engaged vs. Disengaged), Matcher Type (Critical vs. Control), and Trial (whether it was the 1st, 2nd, 3rd, or 4th time that the speaker had described that specific monster).

Table 3 shows the full model output. There was a significant main effect of Trial, such that description length decreased with repeated descriptions (β = -0.14, $p < .001$). There was also a significant main effect of Matcher Type: speakers used longer descriptions when talking to a critical matcher compared to a control matcher (β = -0.54, $p < .001$). Condition was significant as well, with speakers using longer descriptions in the Disengaged condition compared to the Engaged condition (β = 0.59, $p < .01$). There was a significant interaction between Trial and Matcher Type, such that the reduction in description length with repeated reference was greater for critical than control Matchers (β = 0.09, $p < .05$). Critically, the interaction between Condition and Matcher Type was significant (β = -1.09, $p < .001$); as predicted, speakers used more words when addressing the critical matcher in the Disengaged condition compared to the critical matcher in the Engaged condition ($p < .001$), whereas speakers used a similar number of words when addressing control matchers in the Disengaged and Engaged conditions (p = .77). Figure 8 shows the average description length for each matcher type in each condition, collapsed across the trial order variable (as our primary hypotheses do not concern changes over time).

Table 3

Full mixed effects model for Experiment 2

	Estimate	SE	z-value			Var.	Stand. Dev.
Fixed				*Random*			
Trial	-0.14	0.03	-4.94***	*Subject*	(Intercept)	0.230	0.479
Condition	0.59	0.18	3.25**		Trial	0.017	0.132
Matcher Type	-0.54	0.08	-6.82***		Matcher Type	0.152	0.390
Trial : Condition	-0.06	0.06	-1.00		Trial : Matcher Type	0.013	0.115
Trial : Matcher Type	0.09	0.04	2.03*	*Item*	(Intercept)	0.042	0.426
Condition : Matcher Type	-1.09	0.16	-7.00***		Condition	0.046	0.214
Trial : Condition : Matcher Type	-0.04	0.08	-0.48				

Note: * = significant at $p < .05$, ** = significant at $p < .01$, *** = significant at $p < .001$.

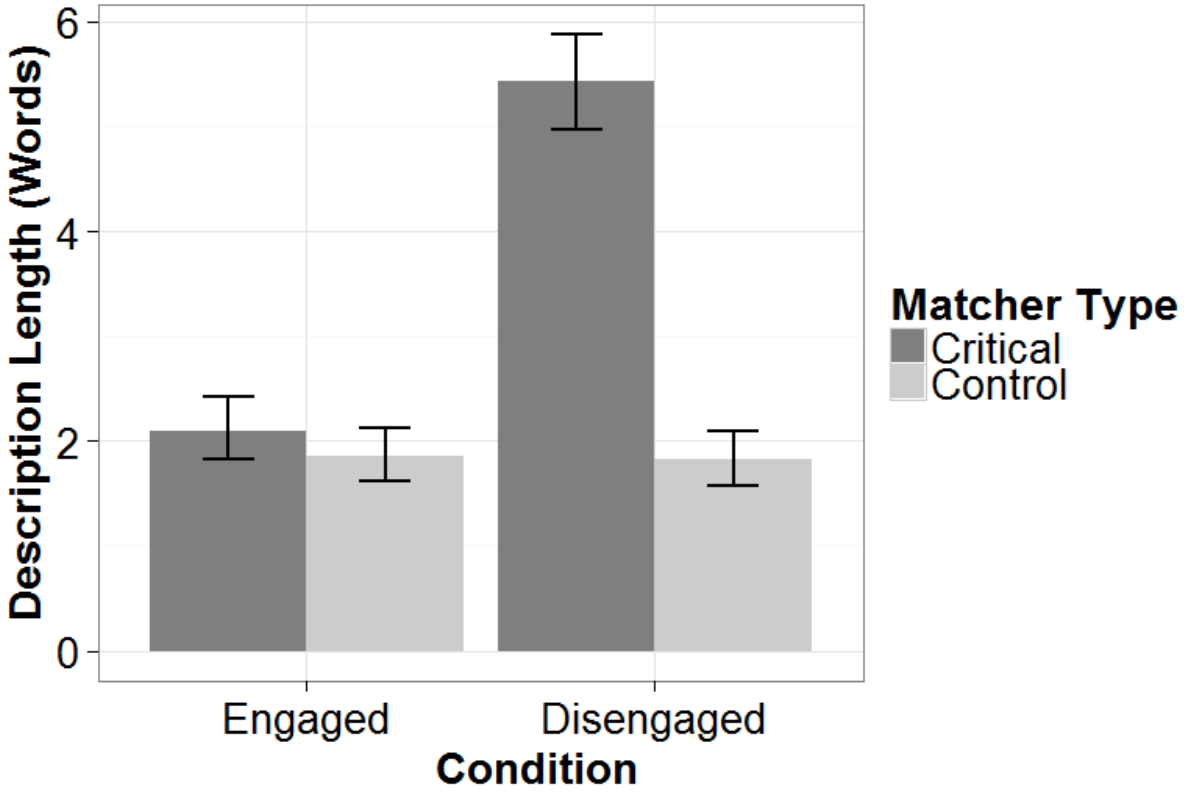


Figure 8. The mean description length for utterances in each condition, broken down by which matcher the utterance was addressed to (data are collapsed across trial). In the Disengaged condition, the critical matcher is the one who was disengaged during the name learning. In the Engaged condition, one matcher was randomly assigned to the role of critical matcher, and their behavior did not differ from the control matcher. Error bars indicate standard error.

3.3 Discussion

In Experiment 2, speakers learned names of monsters with either two fully engaged matchers, or one disengaged matcher and one engaged matcher, then, they completed a referential communication task in which they alternated between addressing the two

matchers. Consistent with our predictions, we found that speakers used longer descriptions when talking to previously disengaged matchers compared to previously engaged matchers. This finding suggests that speakers take into consideration partner attentiveness when assessing whether what has been said is or is not common ground with that partner; these representations of common ground in turn affect subsequent audience design for that particular partner.

Unlike Experiment 1, speakers designed these utterances in a partner specific manner: When one of the two lab assistants had previously been disengaged, speakers produced longer utterances for that previously disengaged matcher, but not for the other matcher. Why were speakers partner-specific in Experiment 2 but not in Experiment 1? The fact that the disengaged behavior in Experiment 1 happened during a task unrelated to the critical test likely played a key role. In Experiment 3, we directly test this interpretation of our findings by manipulating when during the interaction the disengaged behavior takes place.

CHAPTER 4

EXPERIMENT 3

The aim of Experiment 3 was to conceptually replicate Experiments 1-2 in a single experiment in order to test the hypothesis that addressee engagement has specific, rather than generic, effects on assumptions about what information is common ground. The design of Experiment 3 was similar to that of Experiment 2 but with the addition of an unrelated communicative game before the name learning task. In the disengaged condition, the critical manipulation was whether the lab assistant's disengaged behavior occurred during the unrelated communicative game, or during the name learning task. We hypothesize that speakers are not only sensitive to the presence of disengaged behaviors when assessing a partner's knowledge, but that the inferences about (the lack of) common ground due to that disengaged behavior are specific to the information exchanged during the period of inattention / lack of engagement. If this hypothesis is correct, speakers should design longer referring expressions for disengaged matchers, but only when that partner was disengaged during the name-learning task. However, if speakers are sensitive to disengaged behavior, but these effects are non-specific, speakers should not take into account the relevance of that behavior to the topic at hand, and should produce longer utterances at test when addressing the disengaged matcher, regardless of whether the disengaged behavior occurred during the irrelevant communication game or during the test-relevant name learning task.

4.1 Methods

Participants

46 native English speaking undergraduates at the University of Illinois Urbana-Champaign area served as participants for partial course credit. An additional 8 participants were run and excluded from analysis for scoring below the accuracy threshold.

Procedure

Like Experiment 2, each participant took part in the task along with two laboratory assistants. The procedure was identical to Experiment 2, except for the addition of a communication game before the name learning phase.

The communication game was similar to the "Heads Up" game: the three partners took turns holding up a card that showed the name of a historical, pop culture, or fictional person on it. The objective of the Heads Up game is for the person holding the card to guess the name on the card based on information provided by the other participants. The person holding the card was allowed to ask yes or no questions. The other participants were not allowed to say the name on the card, but all other responses were allowed. Participants played this game until they completed 33 trials (11 per participant) or for 15 minutes, whichever happened sooner.

Participants were randomly assigned to one of three between-subjects conditions. In the Engaged condition (n=14), both matchers were engaged throughout the entire experiment. In the Disengaged-Game condition (n=16), the critical matcher was disengaged during the Heads Up game, but was engaged for all other tasks; in this condition the control matcher was engaged throughout the entire study. Finally, in the

Disengaged-Name condition (n=17), the critical matcher was disengaged during the name learning task, but was engaged for all other tasks; in this condition the control matcher was engaged throughout the study. The pairings of names and monster illustrations was counterbalanced across 4 different experimental lists; each participant was assigned to a single list. Overall, the whole task took forty minutes.

Predictions

If assessment of addressee knowledge is sensitive to addressee engagement during the specific circumstances during which the relevant information was communicated, speakers should produce long descriptions for the disengaged addressee in the Disengaged-Name condition, whereas short descriptions should be produced for the disengaged addressee in the Disengaged-Game condition, and for all engaged addressees. However, if addressee engagement has non-specific effects, speakers should assume a lack of common ground for the disengaged addressee, regardless of when the disengaged behavior occurred. If so, this predicts that the speaker should produce longer descriptions for disengaged than engaged matchers, in both the Disengaged-Name and Disengaged-Game conditions.

4.2 Results

Data Exclusions

The exclusion criteria were identical to Experiment 2, so only the speaker's utterances were analyzed and trials where the speaker used the incorrect name at test were excluded, eliminating 2.4% of the trials. The average accuracy for the remaining 46 participants was 97.9%.

Analysis

The speakers' descriptions of the monsters during the sorting task were coded for the number of words in the same way as in Experiments 1-2. The data were analyzed in the same way as Experiment 2, using a mixed effect linear model with a Poisson distribution. The predictor variables in the model were Condition (Engaged, Disengaged-Game, Disengaged-Name), Matcher Type (Critical vs. Control), and Trial (whether it was the 1st, 2nd, 3rd, or 4th time that the speaker had described that specific monster). Condition was Helmert coded to first compare the Engaged condition to both Disengaged conditions, and then to compare the Disengaged-Game condition to the Disengaged-Name condition directly. As before, Matcher Type indexed the critical and control matchers. In the two Disengaged conditions, the critical matcher was the matcher who showed disengaged behavior, and the control matcher was the matcher who always showed engaged behavior. In the Engaged condition, one matcher was randomly assigned to each role (since both matchers were engaged in this condition).

Table 4 shows the full output of the model. There was a significant main effect of Trial, such that descriptions became shorter over time with repeated reference (β = -0.11, p < .001). A main effect of Matcher Type (β = 0.12, p < .01) was qualified by the predicted interaction between Condition and Matcher (β = -0.29, p < .001): Whereas speakers used more words when addressing the critical than the control matcher in the Disengaged-Name condition (p < .001), by contrast, in the Disengaged-Game condition there was no effect of matcher type (p = .65). Figure 9 shows the average description length for each matcher type in each condition, collapsed across trial (there were no specific hypotheses related to the trial variable).

Table 4

Full mixed effects model for Experiment 3

	Estimate	SE	z-value			Var.	Stand. Dev.
Fixed				*Random*			
Trial	-0.11	0.02	-4.39***	*Subject*	(Intercept)	0.24	0.49
Condition 1 (Engaged vs. Disengaged)	0.021	0.16	0.90		Trial	0.02	0.13
Condition 2 (Disengaged-Game vs. Disengaged-Name)	0.08	0.18	0.46	*Item*	(Intercept)	0.02	0.14
Matcher Type	-0.09	0.03	-2.66**		Condition 1	0.01	0.11
Trial : Condition 1	0.02	0.05	0.38		Condition 2	0.03	0.16
Trial : Condition 2	0.06	0.06	1.11				
Condition 1 : Matcher Type	-0.12	0.07	-1.61				
Condition 2 : Matcher Type	-0.29	0.08	-3.63***				
Trial : Condition 1 : Matcher Type	-0.03	0.07	-0.49				
Trial : Condition 2 : Matcher Type	-0.08	0.07	-1.08				

Note: * = significant at p < .05, ** = significant at p < .01, *** = significant at p < .001.

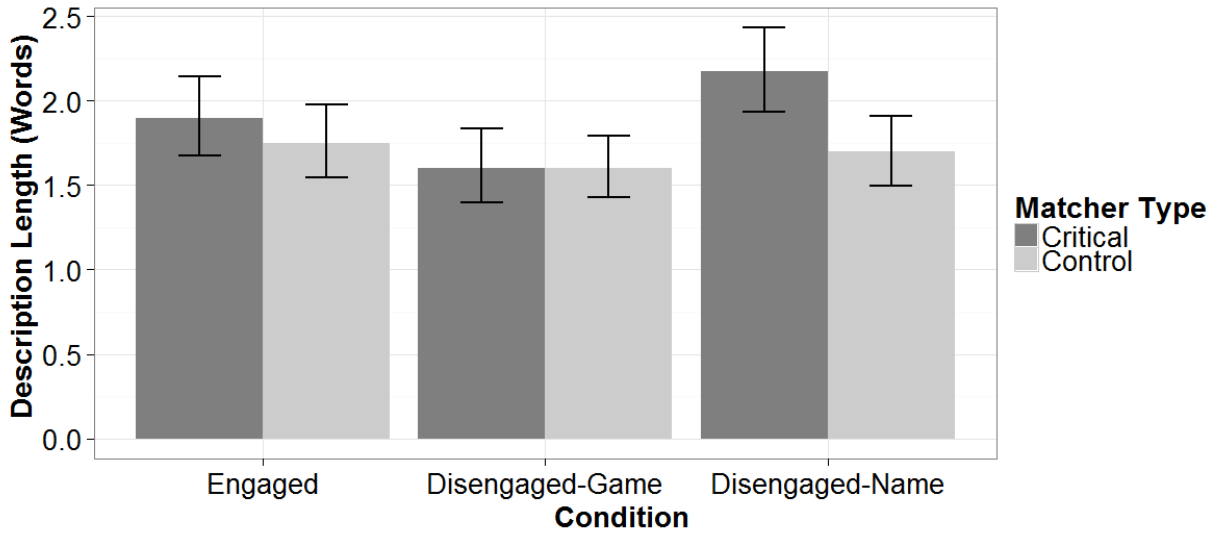


Figure 9. The mean description length for utterances in each condition, broken down by which matcher the utterance was addressed to, and collapsed across trial. In the Disengaged conditions, the critical matcher is the one who was disengaged (either during the name learning or the game). In the Engaged condition, one matcher was randomly assigned to the role of critical matcher, and their behavior did not differ from the control matcher. Error bars indicate standard error.

4.3 Discussion

Speakers in Experiment 3 produced longer descriptions for the disengaged matcher, but only in the Disengaged-Name condition, demonstrating that the inattention of a discourse partner has specific effects on a conversational partner's inferences about what is and is not common ground. We note that the magnitude of the engagement effect is smaller in the Disengaged-Name condition of Experiment 3, compared to the effect observed in Experiment 2. While the name-learning and referential communication tasks were identical in Experiments 2 and 3, all participants in Experiment 3 additionally completed the Heads Up game prior to name learning. Thus we suspect that the smaller effect in Experiment 3

may owe to participant fatigue. That said, the findings in the Disengaged-Name condition of Experiment 3 conceptually replicate and extend the findings of Experiment 2, demonstrating the robustness of this effect.

CHAPTER 5

GENERAL DISCUSSION

Three experiments examined the relationship between the engagement of a conversational partner and inferences about common ground. According to the classic view, conversational participants use simple heuristics to make inferences about what is and is not common ground; a speaker can assume common ground for what they have said when the (rational) listener is co-present and attentive (Clark & Marshall 1978; 1981). These representations of common ground, in turn, guide language use in conversational settings, with speakers providing longer referential descriptions when common ground cannot be assumed (Wilkes-Gibbs & Clark, 1992; Yoon & Brown-Schmidt, 2014). Here, we tested the hypothesis that in conversation, speakers only assume common ground for what they have said when the co-present addressee is attentive in the conversation.

The present findings support this hypothesis, and in addition, show that these effects of addressee attentiveness are specific to the information exchanged during the period of inattention. In Experiment 1, speakers who had interacted with a partner who was disengaged in a prior, irrelevant, task did not provide more details when addressing the disengaged partner than the engaged partner, suggesting that simply exhibiting inattention at some point in the conversation does not cause the speaker to assume that the addressee would fail to understand later communicated information.

To test this interpretation of the data, in Experiment 2, the disengaged behavior

occurred in a prior task that was relevant to the critical referential communication task; consistent with our predictions, speakers used longer descriptions when talking to previously disengaged partners compared to previously engaged partners. Experiment 3 directly tested the specificity hypothesis by manipulating when the disengaged behavior happened, and found that speakers were only more descriptive when talking to partners that were disengaged when they were learning the information that was relevant to success in the current task.

Viewed from the framework of least collaborative effort (Clark & Wilkes-Gibbs, 1986) speakers only generate longer, more effortful descriptions when there is direct evidence that the added cost of producing a long description is necessary to achieve successful communication. When a partner was disengaged during a part of the conversation that was irrelevant to success in the referential communication task (the map task in Experiment 1; the Heads-Up game in the Disengaged-Game condition of Experiment 3), speakers produced similar descriptions for engaged and disengaged matchers, suggesting that both were assumed to have learned the names of the monsters. It was only when the partner exhibited a lack of attention during the relevant, name learning task (Experiment 2 and Disengaged-Name condition of Experiment 3) that speakers produced longer descriptions of the monsters for the disengaged matcher. Thus speakers appear to be sensitive to the fact that attention in a conversation is not static, and can fluctuate over time. Speakers use changes in a partner's level of engagement to guide inferences about what parts of the current experience made it into the common ground, and what parts were missed. Unlike the flat-out denials that were necessary to prevent formation of common ground in previous work (Brown-Schmidt, 2009), the cues used in

the present work were fairly subtle and mimic the kinds of disengagement often encountered in the real world. Furthermore, participants were given no advance warning about the behavior of their matchers, meaning these behaviors were salient enough to be observed without needing to draw attention to them.

The audience design effects in the Disengaged condition of Experiment 2 and the Disengaged-Name condition of Experiment 3 are noteworthy because success in these conditions required the speaker to randomly alternate, from trial to trial which matcher they were addressing. In prior studies of three-party conversation that involved a similar random alternation, speakers successfully produced long descriptions for a matcher who was completely naïve about the names of the referenced stimuli, and short descriptions for a matcher who was knowledgeable (Yoon & Brown-Schmidt, submitted). The present study was more challenging, however, because both matchers were familiar with the monster stimuli -- all three partners participated in the name-learning session together. Such situations in which different partners are associated with similar experiences are known to pose challenges for audience design processes (Horton & Gerrig, 2005), making it all the more impressive that speakers were successful here. An open question is how the distinct interactive styles of the two matchers influence subsequent representations of the conversation. In their study of the role of listener inattentiveness on storytelling, Pasupathi, et al. (1998) brought speakers back to the lab after a 3-week delay to test story recall. They found that speakers who had told stories to inattentive listeners recalled less about those stories compared to speakers who addressed attentive listeners. Taken together with our own findings, the emerging picture is that the process of interacting with disengaged addressees may not only affect conversational processes in the moment (such as whether to

assume common ground, or whether to elaborate a referential description), but in addition, may affect a speaker's memorial representation of that interaction well after the fact.

A final consideration is the specific locus of the inattention effect. Recall that in prior studies of the role of common ground in on-line language comprehension found that common ground was not assumed when a conversational partner had previously denied what had been said (Brown-Schmidt, 2009), but that varying degrees of feedback hypothesized to provide different amounts of evidence for common ground did not have substantively different effects on common ground (Brown-Schmidt, 2012; Brown-Schmidt & Fraundorf, 2015). Yet other work shows that even minimal behaviors such as head nods (Matarazzo, et al., 1964), can influence what speakers say in conversation. Here we alternate between the terms (dis)engagement and (in)attention, and indeed our engaged and disengaged matchers differed in a host of physical, verbal and social behaviors that indicate the presence or absence of engagement and attention. This simultaneous manipulation of a large number of interactional variables was an intentional attempt to mimic the every-day experience of conversing with a partner who is attentive or one who is distracted (e.g., commonly these days, by a smartphone). Whether one, some, or all of these behaviors are necessary in order to elicit the observed effects of attention/engagement remains an open question for future work.

In conclusion, the process by which conversational partners form common ground for what they have been discussing takes into account information about jointly experienced events and the partner's attention to those events. While simultaneously attending to a given physical or linguistic event cannot guarantee mutual knowledge for that event, speakers use the heuristic of simultaneity of attention to infer what is common

ground with the listener. To return to an earlier example, if your friend pulls out her phone in the middle of your conversation, you are likely to assume that she missed what you were saying at that point in time alone. Our findings offer clear support for Clark and Marshall's (1978; 1981) proposal that the simultaneity of attention supports assumptions about common ground, and further, show that inferences regarding the attentiveness of a conversational partner are specific to information that is relevant to what was (or was not) attended to in the conversation.

REFERENCES

Anderson, A. H., Bader, M., Bard, E. G., Boyle, E., Doherty, G., Garrod, S., ... & Sotillo, C. (1991). The HCRC map task corpus. *Language and speech*, *34*(4), 351-366.

Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language, 68*, 255–278.

Bavelas, J. B., Coates, L., & Johnson, T. (2000). Listeners as co-narrators. *Journal of personality and social psychology*, *79*(6), 941.

Bertsch, S., Pesta, B. J., Wiscott, R., & McDaniel, M. A. (2007). The generation effect: A meta-analytic review. *Memory & cognition*, *35*(2), 201-210.

Brennan, S. E., & Clark, H. H. (1996). Conceptual pacts and lexical choice in conversation. *Journal of Experimental Psychology: Learning, Memory & Cognition, 22*, 482–493.

Brown, G., Anderson, A., Yule, G., and Shillcock, R. (1983). *Teaching Talk*. Cambridge, U.K.: Cambridge University Press.

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

Brown-Schmidt, S. (2012). Beyond common and privileged: gradient representations of common ground in real-time language use. *Language and Cognitive Processes, 27*, 62–89.

Brown-Schmidt, S., & Fraundorf, S. H. (2015). Interpretation of informational questions modulated by joint knowledge and intonational contours. *Journal of Memory and Language, 84*, 49-74.

Brown-Schmidt, S., Gunlogson, C., & Tanenhaus, M. K. (2008). Addressees distinguish shared from private information when interpreting questions during interactive conversation. *Cognition, 107*, 1122–1134.

Brown-Schmidt, S., Yoon, S. O., & Ryskin, R. A. (2015). People as Contexts in Conversation. *Psychology of Learning and Motivation*, *62*, 59-99.

Clark, H. H. (1996). *Using language*. Cambridge University Press.

Clark, H. H., & Bangerter, A. (2004). Changing ideas about reference. In *Experimental pragmatics* (pp. 25-49). Palgrave Macmillan UK.

Clark, H. H., & Krych, M. A. (2004). Speaking while monitoring addressees for understanding. *Journal of Memory and Language, 50*, 62–81.

Clark, H. H., & Marshall, C. R. (1978). Reference diaries. In D. L. Waltz (Ed.), *TINLAP-2: Theoretical Issues in Natural language Processing-2. Association for computing Machinery, New York* (pp. 57–63).

Clark, H. H., & Marshall, C. M. (1981). Definite reference and mutual knowledge. In A. K. Joshi, B. L. Webber, & I. A. Sag (Eds.), *Elements of discourse understanding* (pp. 10–63). Cambridge: Cambridge University Press.

Clark, H. H., & Murphy, G. L. (1982). Audience design in meaning and reference. *Advances in psychology*, *9*, 287-299.

Clark, H. H., & Wilkes-Gibbs, D. (1986). Referring as a collaborative process. *Cognition, 22*, 1–39.

Fussell, S. R., & Krauss, R. M. (1989). The effects of intended audience on message production and comprehension: Reference in a common ground framework. *Journal of Experimental Social Psychology*, *25*(3), 203-219.

Galati, A., & Brennan, S. E. (2010). Attenuating information in spoken communication: for the speaker, or for the addressee? *Journal of Memory and Language, 62*, 35–51.

Gorman, K. S., Gegg-Harrison, W. M., Marshall, C. R., and Tanenhaus, M. K. (2013). What's learned together stays together: speakers' choice of referring expression reflects shared experience. *Journal of Experimental Psychology: Learning, Memory & Cognition, 39*, 843–854.

Heller, D., Gorman, K. S., & Tanenhaus, M. K. (2012). To name or to describe: shared knowledge affects referential form. *Topics in cognitive science, 4*(2), 290-305.

Holler, J., & Wilkin, K. (2009). Communicating common ground: how mutually shared knowledge influences the representation of semantic information in speech and gesture in a narrative task. *Language and Cognitive Processes, 24*, 267-289.

Horton, W. S., & Gerrig, R. J. (2002). Speakers' experiences and audience design: knowing when and how to adjust utterances to addressees. *Journal of Memory and Language, 47*, 589–606.

Horton, W. S., & Gerrig, R. J. (2005). The impact of memory demands on audience design during language production. *Cognition, 96*, 127–142.

Horton, W. S., & Spieler, D. H. (2007). Age-related differences in communication and audience design. *Psychology and Aging, 22*, 281–290.

Isaacs, E., & Clark, H. H. (1987). References in conversation between experts and novices. *Journal of Experimental Psychology: General, 116*, 26–37.

Krauss, R, M., & Weinheimer, S. (1964). Changes in reference phrases as a function of frequency of usage in social interaction. A preliminary study. *Psychonomic Science, 1*, 113-114.

Krauss, R. M., & Weinheimer, S. (1966). Concurrent feedback, confirmation, and the encoding of referents in verbal communication. *Journal of Personality and Social Psychology*, *4*, 343–346.

Kuhlen, A. K., & Brennan, S. E. (2010). Anticipating distracted addressees: How speakers' expectations and addressees' feedback influence storytelling. *Discourse Processes*, *47*(7), 567-587.

Matarazzo, J. D., Saslow, G., Wiens, A. N., Weitman, M., & Allen, B. V. (1982). Interviewer head nodding and interviewee speech durations. *Psychotherapy, 1*, 54-63.

Pasupathi, M., Stallworth, L. M., & Murdoch, K. (1998). How what we tell becomes what we know: Listener effects on speakers' long-term memory for events. *Discourse Processes*, *26*(1), 1-25.

Roque, A., & Traum, D. (2008, June). Degrees of grounding based on evidence of understanding. In *Proceedings of the 9th SIGdial Workshop on Discourse and Dialogue* (pp. 54-63). Association for Computational Linguistics.

Schober, M. F., & Clark, H. H. (1989). Understanding by addressees and overhearers. *Cognitive Psychology, 21*, 211–232.

Slamecka, N. J., & Graf, P. (1978). The generation effect: Delineation of a phenomenon. *Journal of experimental Psychology: Human learning and Memory*, *4*(6), 592.

Wilkes-Gibbs, D., & Clark, H. H. (1992). Coordinating beliefs in conversation. *Journal of Memory and Language, 31*, 183–194.

Yoon, S. O., & Brown-Schmidt, S. (submitted). Aim low: Mechanisms of audience design in multiparty conversation.

Yoon, S.O. & Brown-Schmidt, S. (2014). Adjusting conceptual pacts in three-party conversation. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 40*, 919-937.