ELSEVIER

Cognition 59 (1996) 91–117

COGNITION

When do speakers take into account common ground?

William S. Horton, Boaz Keysar*

Department of Psychology, University of Chicago, 5848 South University Avenue, Chicago, IL 60637, USA

Received 24 October 1994, final version accepted 16 August 1995

Abstract

What role does common ground play in the production of utterances? We outline and test two models. One model assumes that common ground is involved in initial utterance planning, while the other model assumes that it only plays a role in monitoring. To compare these models, we focus on common ground as evidenced in physical co-presence. We had speakers describe objects for listeners in a modified version of the referential communication task. While descriptions under no time constraints appeared to incorporate common ground with the listener, common ground was not used when the speakers were under time pressure. These results suggest that speakers do not engage in audience design in the initial planning of utterances; instead, they monitor those plans for violations of common ground.

1. Introduction

Imagine the hypothetical question posed by speaker A to listener B: "Did you eat of the tree from which I had forbidden you to eat?" The production of this utterance required the operation of a variety of processes, from message planning through formulation of the utterance to its actual articulation (e.g., Bock, 1995; Garrett, 1988; Levelt, 1989). As part of this complex production process, the speaker applied linguistic knowledge of form and meaning via a variety of subprocesses. Researchers in the area have studied extensively the operation of many of these subprocesses, including the selection and retrieval of words from the lexicon (e.g., Butterworth, 1989; Dell & O'Seaghdha, 1991; Harley, 1993), the construction of syntactic

* Corresponding author. Fax: 312 702 0886; e-mail: boaz@ccp.uchicago.edu.

SSDI 0010-0277(95)0690-7

structures (e.g., Bock, 1986, 1990; Bock & Kroch, 1989; Bock, Loebell, & Morey, 1992; Cooper & Paccia-Cooper, 1980) and prosodic structures (e.g., Cutler, 1980; Cutler & Isard, 1980; Ferreira, 1993; Meyer, 1994). Utterance production also requires a relatively understudied facet of the process: the way speakers take their listeners' perspective into account. What role does knowledge about the listener play in the production process? In the forbidden fruit example, the speaker seems to be sensitive to the listener's knowledge about possible tree referents. In this utterance the speaker relies on mutual knowledge: the speaker had forbidden the listener from eating of one type of tree, and this fact is mutually known to them (Clark & Marshall, 1981; Clark, Schreuder, & Buttrick, 1983). This knowledge allows the speaker to contrast the tree in question with all other trees, assuming the listener will use the same information to understand the utterance. The focus of our paper is on the way such knowledge of common ground or mutual knowledge is applied in the construction of utterances.

2. Referential communication and common ground

Much of communication is referential – where a speaker attempts to identify for an addressee a particular referent in the world. When the speaker asks about "the tree" he or she intends the listener to pick out a particular type of tree in the world. How do speakers construct such referential descriptions? Olson (1970) proposed that the nature of the description depends on the particular contrast set. Given that the goal of the speaker is to identify the referent uniquely, the description must pick out the unique referent and only that referent. Had the speaker forbidden the listener to eat from more than one tree, then the description "the tree from which I had forbidden you to eat" would be inadequate.

Clark and his colleagues argued that the guiding principle in the construction of referential descriptions is not merely the particular contrast set but the fact that it constitutes the common ground between the speaker and the addressee (e.g., Clark & Haviland, 1977; Clark & Marshall, 1981; Clark et al., 1983). Common ground is information which is mutually believed by both parties. When speakers rely on common ground with their addressee, they are being cooperative (Grice, 1975). Under this conception, then, even when the speaker had forbidden two kinds of trees, the description might still be adequate if there is an additional difference between them which is mutually salient to the parties. Indeed, Clark et al. (1983) demonstrated that salience affects speakers' referential descriptions. In addition, a variety of studies demonstrate that once common ground is established it is used to guide speakers' selection of referring expressions, allowing the shortening of these expressions and establishing the conditions for a transition from descriptions to names (e.g., Clark and Wilkes-Gibbs, 1986; Glucksberg, Krauss, & Weisberg, 1966; Isaacs & Clark, 1987; Krauss, 1987). In addition,

developmental studies reveal that en route to becoming competent adult speakers, children acquire the ability to use common ground in referential communication (Ackerman, Szymanski, & Silver, 1990; Deutsch & Pechmann, 1982). In general, then, referential descriptions follow principles of "audience design" (Clark & Murphy, 1982). The speaker is said to adhere to the *Principle of Optimal Design*, in which "the speaker intends each addressee to base his inferences not on just *any* knowledge or beliefs he may have, but only on their *mutual* knowledge or beliefs – their common ground" (Clark, 1992, p. 81, emphases in the original).

This body of work suggests that common ground affects the nature of referential descriptions. Because these studies look at the final product of the production system, they do not reveal the role of common ground information in the production process. For example, we do not know at what point in utterance construction the production system incorporates the speaker's knowledge of common ground. We do not know whether common ground guides the initial planning of utterances or whether the incorporation of common ground occurs rather late in the process. The study presented here investigates these issues directly.

3. Models of common ground in utterance production

3.1. Components of utterance production

In order to identify the role of common ground in production, we will first outline some of the characteristics of the production system. While the exact nature of the elements of the system is controversial, and the extent to which the system is interactive or modular is under debate (e.g., Dell, 1986; Dell & O'Seaghdha, 1992; Garrett 1976; Schriefers, Meyer, & Levelt, 1990), there is general agreement about the role of several functional components (e.g., Bock, 1995; Garrett, 1988; Fromkin, 1971; Levelt, 1989). As Levelt (1989) suggests in his "blueprint for the speaker" (p. 9), the production of an utterance requires an initial preverbal intention (but see the debate between Butterworth & Hadar, 1989, and McNeill, 1989). The preverbal message is put into a linguistic form via grammatical and phonological encoding. This results in a phonetic plan which is converted to an articulation plan. The speaker also monitors the production system to detect and intercept a variety of possible errors. For example, Levelt (1983) presents evidence for monitoring of the phonetic plan of utterances as well as monitoring of overt speech.

In general, then, utterance production involves both planning aspects of the utterances and the monitoring of these plans (Laver, 1980; De Smedt & Kempen, 1987). The goal of monitoring is to shape the final product into a communicative utterance. Speakers monitor for linguistic well-formedness and for the extent to which the linguistic form reflects their intention. For

example, speakers monitor for word choice and for pragmatic adequacy of utterances (e.g., Levelt, 1983). Such monitoring might occur after an error was produced, possibly resulting in an overt repair, or it might occur earlier in the process, resulting in an internal revision of the utterance prior to articulation. Given these two general elements of the production process – planning and error correction – what role does the speaker's common ground with the addressee play in utterance production? We will describe and evaluate two general models that differ with respect to the use of common ground in utterance planning and the nature of monitoring for errors *vis-à-vis* common ground.

3.2. *The "Initial Design" model*

According to this model, speakers apply the Principle of Optimal Design in their utterance plan. The plan takes the addressee into account, and incorporates only information which is part of the common ground between the speaker and the addressee. According to the optimal design principle, the speaker restricts herself to the common ground information because it is shared, and because deviating from this information might result in an infelicitous utterance. The utterance plan, then, is tailored to the addressee's perspective from the outset. This leaves relatively little work for the monitoring process with respect to common ground. Occasionally, a nonsystematic error might result in an utterance which is inadequate *vis-à-vis* the common ground with the addressee. According to this model, the role of monitoring is only to detect such nonsystematic errors.

To illustrate the operation of the Initial Design model consider the referential descriptions in the following example which uses a contrast set. According to Clark et al. (1983), interlocutors do not simply use a contrast set to interpret a reference, but they use the contrast set which is part of their common ground. Suppose that a bakery has two types of bread, both of which are round but one of them is about 5 inches and the other about 7 inches in diameter. You are interested in the 7-inch loaf and you say to the baker, "I'd like a large loaf of bread, please." Given that a 7-inch bread is not large in any "objective" sense, and that it is not even large for a round loaf of bread, how would the baker identify the referent of your description? The baker will probably interpret the adjective "large" in the context of other breads which are present. Given that the 7-inch bread is large compared to the 5-inch bread she will have no difficulty identifying the intended referent. What is important for our discussion is that not only is the contrast between the two diameters available for the baker, it is also mutually known to you, the speaker, and to the baker. This contrast is mutually known because it is physically co-present (Clark & Marshall, 1981). According to the Initial Design model, *because* the contrast set is part of the common ground, you used this information in planning your utterance. Common ground, then, contributed to the very planning of the

utterance. Given such planning which adequately incorporated common ground information, the monitoring function was not needed.

The Initial Design model makes sense – why not incorporate into the initial plan common ground information if it is readily available? Just like incorporating any other utterance-relevant information into the plan, common ground could be used to tailor the plan to the addressee's knowledge. Yet, an alternative model might also be plausible for the following reason: knowledge about what the other knows and does not know might be too costly to use routinely in utterance planning. In fact, in many cases it might not be necessary to use common ground to tailor utterances to addressees. Often, what is available to speakers might also happen to be part of common ground. In these cases, by using what is available to them, speakers would also be using information which is part of common ground. The Monitoring and Adjustment model incorporates these assumptions.

3.3. The "*Monitoring and Adjustment*" model

In contrast to the Initial Design model, the Monitoring and Adjustment model assumes that common ground does not play a role in the initial plan of utterances: the initial plan is not designed for the specific knowledge of the addressee. Speakers plan their utterances using information which is available to them regardless of whether or not the information they use is part of the common ground with the addressee. Consequently, the plan might occasionally rely on information which is inaccessible to the addressee, which might result in an utterance that is not fully communicative. Alternatively, the plan might not make sufficient use of common ground in order to produce a properly tailored utterance. The model assumes that speakers monitor their productions, intercepting plans that rely on non-shared information and those that do not rely sufficiently on shared information. Such plans are then revised to accommodate common ground. According to this model, then, common ground operates as part of a correction mechanism that is part of the monitoring function of production.

Consider the way the Monitoring and Adjustment model would account for the large loaf of bread example. Recall that according to the Initial Design model your utterance plan incorporated the fact that the contrast set was co-present. In contrast, according to the Monitoring and Adjustment model, you used the referring expression not because the contrast was part of common ground but because it was salient for you. The reason it was salient is that you were looking right at the two kinds of breads. In other words, your utterance plan relied on the size contrast because is was present, *not* co-present. The fact that it is also co-present and hence mutually known was not considered in the plan. In this example, both models suggest the use of the contrast set information. Unlike the Initial Design model, the Monitoring and Adjustment model suggests that in-

formation that is present for the speaker would be used regardless of whether or not it is also co-present with the listener. If the information also happens to be co-present or shared then monitoring would not be required, but if it is not co-present then the utterance plan might be monitored and revised.

3.4. Differential predictions of the models

The main difference between the two models concerns the role of common ground during initial planning. The Initial Design model suggests that if one were able to tap the initial plan, that plan would exhibit the incorporation of common ground information: it would rely on shared information. In contrast, the Monitoring and Adjustment model suggests that were we able to uncover the initial plan, it would turn out to be agnostic *vis-à-vis* common ground: it would exhibit a reliance on information regardless of whether or not the information is shared with the addressee. Though both models predict that the end-product of the process, the utterance, would rely on common ground information, they predict this for different reasons. According to the Monitoring and Adjustment model, the utterance relies on common ground because the monitoring process corrects for violations of common ground. According to the Initial Design model, the utterance relies on common ground because it was so planned from the outset. The critical test between the two models is whether or not the initial plan distinguishes between information which is part of common ground and information which is *not* part of the common ground – that is, privileged information available solely to the speaker. If it does not distinguish between the two types of information, this would support the Monitoring and Adjustment model because it would demonstrate that the initial plan is not restricted to common ground information; instead, it is using information because of its availability to the speaker, not because it is shared with the addressee. In order to differentiate the models, we must attempt to tap the pre-monitoring utterance plan as well as post-monitoring production and compare the extent to which speakers rely on shared and on privileged information.

4. The experimental paradigm and general predictions

To evaluate the models, one must first select a particular instantiation of common ground. One possibility would have been to look at the definiteness of speaker's descriptions (see Clark and Marshall's, 1981, theoretical treatment). Alternatively, one can look at the effect of mutual context on the nature of speaker's descriptions, such as their use of adjectives. Our preliminary investigations revealed that adjective use would be a much more fruitful avenue to pursue because it was more sensitive a measure than

article use. Consequently, we selected an instantiation of common ground suggested by Clark et al.'s (1983) criticism of Olson's (1970) theory of the role of contrast in referential descriptions. They suggest that interlocutors do not simply use a contrast set to form and interpret referential descriptions, but that they use the contrast set which is part of their common ground. A direct way to evaluate the contrast set which interlocutors use is to consider the kind of adjectives they include in their referential descriptions (e.g., "small" vs. "large"). Therefore, we looked at the nature of speakers' descriptive noun phrases as a function of the contrast set available to them.

We investigated speaker's referential descriptions by modifying the referential communication task (e.g., Glucksberg, Krauss, & Higgins, 1975; Krauss & Glucksberg, 1977; Isaacs & Clark, 1987; Schober & Clark, 1989). Subjects were speakers in a communication game. They saw one-half of a computer screen which was separated by a barrier from their addressee who could only see the other half of the screen (Fig. 1a). In each trial, two

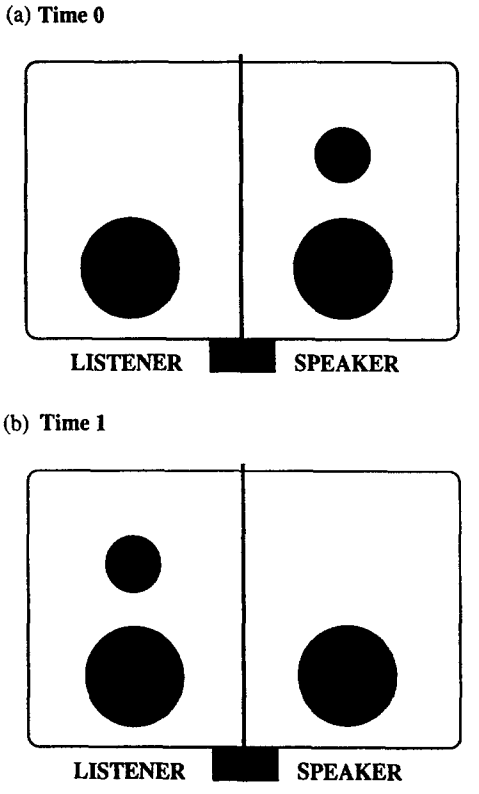

Fig. 1. An example of a trial in the "shared context" condition. In Time 0 the speaker described the upper object, which moved behind the visual barrier into the listener's part of the screen in Time 1.

objects appeared on the speaker's side and then the top object moved to the addressee's side across the barrier (Fig. 1b). The speakers' task was to describe the moving object so that their addressee could identify it. They mutually knew that the nonmoving object appeared on both sides and therefore constituted common ground by virtue of its co-presence (Clark & Marshall, 1981).

How would the speaker refer to the moving circle in Figure 1(a)? Using common ground, the speaker might use the shared contrast set and refer to the moving circle as "a small circle". The use of the adjective "small" might appear like an obvious instance of the use of common ground because the contrast was shared, but this is not necessarily the case. It is possible that the speaker used the contrast set in this description not because it is common information but because it is part of her knowledge. This is precisely how the Monitoring and Adjustment model would account for such an utterance. According to this model, the initial plan used the contrast set because it was available to the speaker. The utterance is also in line with the common ground information but only because the contrast set happens to be shared. The way to distinguish the accounts of the two models is to contrast the case when the context object is shared (e.g., Fig. 1a and b) with a case when this context is privileged to the speaker (e.g., Fig. 2a and b). If the Initial Design model is correct, then speakers should use that context when it is shared but not when it is privileged. According to the Monitoring and Adjustment model, though, the initial plan would use the context information regardless of whether it is shared or privileged. In general, then, one can argue that speakers use common ground only if they rely on the context information when it is shared but *not* when it is privileged. In the context of our experiment, speakers will be said to make use of common ground to the extent that their descriptions rely on the contrast set *less* when the contrast set is privileged than when it is shared.

Suppose that speakers are equally likely to refer to the moving circle as "a small circle" when they know that the context is privileged (Fig. 2a and b) as when it is shared (Fig. 1a and b). We might want to conclude that they are not using common ground because they rely on the context regardless of whether it was shared or not. This conclusion would be misleading, because it is possible that even though the description seems systematically related to the context, the speakers may not have relied on the contrast set (i.e., the nonmoving circle). Instead, they might have thought that the moving circle is small in some absolute sense, or they might have simply believed that such a description is adequate even for someone who does not see the contrast set. To be conservative, then, the crucial test for reliance on context is when the speaker describes the same circle once again but with a different context. Speakers, then, will describe each object a second time; in this case, the same top circle will be presented with a smaller context circle, as in Fig. 3. If the speakers rely on their privileged information in this case, they might refer to the circle as "a large circle". If in this second trial speakers

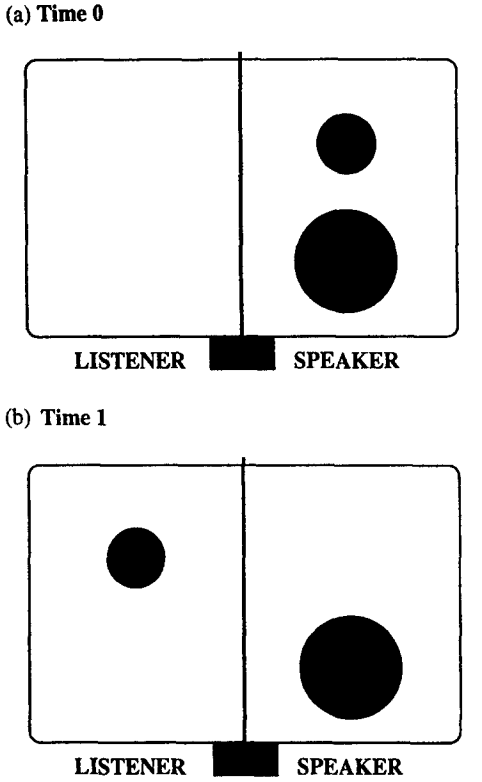

Fig. 2. An example of a trial in the "privileged context" condition. The lower object cannot be seen by the listener and constitutes the speaker's privileged contrast set.

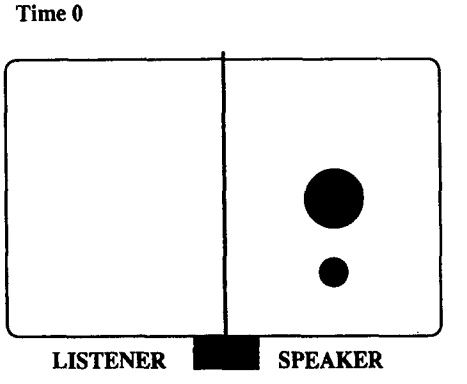

Fig. 3. The second presentation of the item from Fig. 2(a). The same medium-size circle is paired with a smaller circle.

show similar reliance upon the contrast set regardless of whether it is shared or privileged, this would constitute evidence against the role of common ground in utterance planning. In contrast, if the speakers do incorporate common ground into the utterance, they should use referring expressions which make use of shared but not privileged context information. In general, a second description which is systematically related to the speaker's privileged information would suggest that the speaker did not use common ground in the planning of the utterance; a description which does *not* rely on such privileged information could suggest either that the speaker incorporated common ground information in the initial utterance planning (i.e., the Initial Design model) or that the speaker detected use of nonshared information and monitored it out of the final utterance (i.e., the Monitoring and Adjustment model).

To determine whether descriptions which are *not* systematically related to privileged information are the result of initial design or of corrective monitoring, we had some subjects initiate their descriptions quickly. Our assumption is that the pressure to initiate the description would reduce the potential effect of monitoring and adjustment. Consequently, the description should reflect more directly the initial planning of the utterance. Thus, the predictions of the two models are clear: according to the Initial Design model the time pressure manipulation should not systematically alter the nature of the production because the initial plan already uses common ground. Therefore, the second description should rely on shared information but not on privileged information regardless of time pressure. In contrast, according to the Monitoring and Adjustment model the time pressure manipulation should eliminate the correcting effect of monitoring, and speakers should fall back on a plan which does not use common ground. Consequently, time pressure should result in descriptions which are not sensitive to common ground. Therefore, this model predicts that under time pressure speakers' second descriptions would rely on their context to the same degree, regardless of whether it is privileged or shared.

5. Method

5.1. Subjects

Twenty-four native English-speaking undergraduates from the University of Chicago, recruited via flyers posted around campus, participated in this study. Subjects were randomly assigned to an experimental condition and were always given the role of the "speaker". The "listener" was played by a confederate. Data from one subject could not be used due to experimenter error and were replaced with data from an additional subject. All subjects were paid for their participation.

5.2. Apparatus

This experiment was controlled by a Macintosh IIci computer running *Animation Works*, a software package that can present simple animated "movies". The speaker and the listener were separated from one another by a large 1 m × 1.5 m piece of foam board that was extended perpendicular to the computer monitor. This barrier fitted snugly against the monitor, preventing the subjects from seeing any portion of the other side of the screen. The speakers' descriptions were recorded through a clip-on lapel microphone.

5.3. Stimuli

5.3.1. Experimental items

The stimuli were clip-art images of shapes and objects which were relatively common and easily identifiable. We constructed 16 experimental item sets. Each item set had a moving object which was paired with two nonmoving context objects: once with a context object that had a bigger value on a certain dimension (e.g., darker in shade) and once with a context object that had a smaller value on that same dimension (e.g., lighter in shade). Nine of the experimental sets had such "bidirectional" contrast sets, where the salient dimension was altered in both directions. The remaining items consisted mainly of "unidirectional" relationships, in which the relevant dimension was altered in only one of the two items. For example, a moving white star had a black star as its context in one case, making shade the relevant dimension, but it had a white moon as the second context, where shade was no longer diagnostic. With these "unidirectional" items, the contrast dimension was relevant for only one case of each pair and therefore the use of context can be detected only in one case. This did not become apparent until after the data had already been collected. Though analyses with and without the unidirectional items showed identical patterns of results, we focus only on the results of the bidirectional items because they have the potential to be most revealing about reliance on context.

In all experimental item sets the critical object remained unchanged as it moved from one side of the screen to the other. The listener's task was to verify whether or not her moving object matched the speaker's description. Therefore, the experimental items should result in agreement between the speaker and the listener that the object which moved out of speaker's screen was identical to the one which moved into the listener's screen.

5.3.2. Fillers

There were 24 filler sets. In each set the moving object appeared twice, paired with two context objects. Different from the item sets, the objects in the filler items bore no systematic relationship to one another from one presentation to the next. In 20 fillers, the moving object changed as it

moved across the barrier separating the two halves of the screen, usually transforming into some related yet distinct object. Figs. 4(a) and (b) depict one such filler item. The goal of these fillers was to include trials in which the object that moved into the listener's side was different from the one that moved out of the speaker's side of the screen, helping to motivate the experimental task. Also, in order that the experimental items not be confounded with affirmative responses, the moving object stayed constant in four additional filler sets. In total, the moving object stayed constant in half the cases and changed in the other half.

5.3.3. First and second presentations

Each moving object was presented twice: once with one context object, and in the second presentation with the second context object. The experimental sets appeared at fixed, yet randomly determined positions in each presentation. We determined the position of items such that no more than three experimental items appeared in a row. The associated context

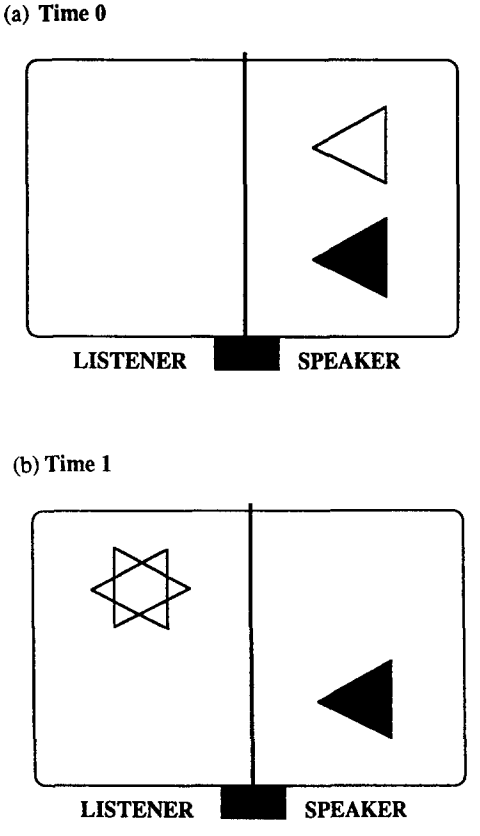

Fig. 4. An example of a negative-trial filler. The object that moved into the listener's screen is different from the one that moved out of the speaker's screen.

objects were randomly assigned to appear in the first or second presentation, with the constraint that the number of context objects with a positive value on the relevant dimension (larger, darker) was equal to the number of objects with a negative value (smaller, lighter) in each presentation. To counterbalance the order of appearance of each context object, two versions of the experiment were constructed by switching the two orders of presentations for half the subjects. This prevented the confounding of a particular contrast set with a particular position. The filler items were randomly placed among the experimental items. The order of the filler items was different in each presentation.

5.4. Procedure

All subjects received the same general introduction to the experiment. They were told that they would be playing the role of the speaker in a communication experiment. After being introduced to the confederate listener, speakers were told that their task was "to describe each object in such a way that will allow the listener to determine whether she sees the same object move". They received a series of examples showing how the objects moved and also how they sometimes changed as they moved across the barrier. The experimenter pointed out the presence of the nonmoving context objects and emphasized that it was not necessary to describe these stationary objects. The experimenter then instructed the listener to reply only "yes" or "no" in response to the speakers' descriptions. The confederate listener participated as if she were really naive to the situation. She listened to the subjects' descriptions, compared them to the moving objects and responded "yes" or "no" based solely on whether or not the descriptions matched the objects. The only explicit directions to the listener as to how to behave came in the four switched-role practice trials to be explained below. Otherwise, she responded directly to the subjects' descriptions.[1]

The experimenter proceeded to describe what the listener would see during the experiment. Half the subjects were told that in addition to the moving objects, the listener would never see any kind of nonmoving context objects (i.e., privileged-context condition); the other half were told that the listener would always see the same context objects (i.e., shared-context condition). The subjects saw examples that illustrated this information in order to ensure that they completely understood the nature of the listener's knowledge about the context objects.

Because the context manipulation was so central to our experiment we wanted to given subjects maximum opportunity to understand the perspective of the listener. As part of the practice phase, the speaker and the listener switched roles for four sample trials. This was primarily important

[1] Subjects were eventually informed that the listener was a confederate. All indicated that they believed that she was an actual subject.

for the privileged-context condition, but all subjects performed the role-switching for uniformity's sake. The experimenter informed the subjects that the purpose of the exercise was to give them "a more complete idea of what the experiment is about" and to allow them to get a feel for the other's point of view. The confederate was instructed beforehand how to describe the moving objects on each of these example trials. After each trial, the experimenter removed the barrier to allow the subject to evaluate the trial. The first role-switching trial contained a moving medium-sized circle, with a stationary larger circle on the confederate's side (and also on the subject's side, if they were in the shared-context condition). The confederate, playing the role of the speaker, described the moving object as "a small circle". The subjects invariably responded "yes", and when the experimenter removed the barrier they had the opportunity to confirm that the trial was successful (i.e., that it was acceptable to reply "yes" to "small circle"). After two fillers, the fourth and final trial once again had the medium-sized circle as before, only this time it was paired with an even smaller context circle. In this case, the confederate described the object as "a large circle", which contradicted her description in the first presentation. When the barrier was removed, subjects had the opportunity to evaluate the description. If they replied "yes", the experimenter reiterated the instructions to describe the objects in the manner most likely to elicit the correct response from the listener. If they said "no", the experimenter pointed out that some descriptions may lead to confusion, and that they should keep in mind that their primary goal is to provide descriptions so that the listener can easily determine whether or not she is seeing the same thing.

The final instructions concerned the speed at which speakers should initiate their descriptions. Half of the subjects were in the "speeded condition". In general, subjects knew which object would be moving because it was always the topmost object, and they were informed that there was a 1.5-second interval between the time the object appeared on the screen and when it began to move. We directed half of the subjects to begin their descriptions no later than the moment when the objects began their movement. If the experimenter felt that a subject was tardy in beginning the descriptions, he reminded the subject once again of the necessity of responding quickly. The remaining subjects were in the "unspeeded" condition. We instructed them to take as much time as they liked in initiating their descriptions.

After receiving all the instructions, subjects had an opportunity to ask questions and then completed five practice trials which were modeled after the actual experiment. When the subjects indicated that they understood the procedure, they clipped on the lapel microphone and the experiment began. The subject initiated each trial by pressing the computer mouse button, which caused the objects to appear in their starting positions. Each trial lasted roughly 10 seconds, after which the screen went blank. The subjects were instructed not to begin the next trial until after both the screen had

gone blank and the listener had responded. Subjects saw all items and fillers in their first presentation. After a 5-minute break, they described the items in their second presentation. The experiment lasted approximately 30 minutes.

5.5. Design

There were two context conditions and two initiation speed conditions. The context was either shared with the listener or privileged to the speaker, and the speaker was either speeded in initiating the utterance or not speeded. In addition to the two experimental manipulations, we separated the data for the first and second presentation of each set because the predictions most directly concern the second presentation, where a description systematically related to the context would more clearly suggest use of that context. This resulted in a 2 (context condition: privileged vs. shared) × 2 (initiation speed: unspeeded vs. speeded) × 2 (presentation: first vs. second) randomized factorial design. Presentation was the only within-subjects factor; the other two were between-subjects.

6. Results

The recordings of the descriptions were transcribed verbatim and coded to reflect reliance on context as we describe below. In general, each speaker seemed to adhere to a particular form across descriptions (e.g., "I see an X moving from my side to yours"). This resulted in little variation in form, as illustrated by the use of articles: in the privileged-context condition, 16% of the descriptions used the definite article, while less than 1% of the descriptions were definite in the shared-context condition (regardless of speed).[2] There was wide variation, however, in length and content. Some descriptions briefly identified the moving object ("An envelope from right to left"), while others (only 5% of the data) gave more elaborate descriptions ("The three quarter of an inch high one and a one one-inch wide white envelope moved from right to left"). The most common type of description was between these two extremes, containing one or two adjectives that modified the noun identifying the critical object ("A small sealed envelope moved right to left"). Given that the extent to which speakers relied on context can be revealed by the kind of adjectives they used to describe the

[2] One might argue that indefinite descriptions cannot be constrastive, but this would be wrong. As our bread example suggests, one could ask for "a large loaf" and would only be understood *in contrast* to a smaller loaf. The prevalence of indefinite descriptions is probably due to examples in the instructions which used the indefinite article, and to the confederate's use of the indefinite description. The issue of definiteness, then, does not bear on our experiment because it was not designed to test this.

moving objects, we coded for the extent to which the adjectives were related to the context.

6.1. Coding

During coding, we were aware of the context object but we were blind to condition – that is, neither did we know if the context was shared or privileged nor if the speaker was speeded or unspeeded. We classified each adjective as either context-related ("small" with a larger envelope) or context-unrelated ("sealed" with a larger sealed envelope) and summed the counts of each type for each condition. Before coding, we spelled out the criteria for a classification scheme that was conservative with respect to classification of an adjective as context-related. Here is a summary of the criteria:

- In general, a context-related adjective was an adjective which specified the value of the target object on its contrasting dimension with the context object. Other adjectives were coded as context-unrelated. Consider a medium gray dinosaur, paired with either a lighter or darker gray dinosaur context object. When paired with the lighter dinosaur, we coded as context-related the adjective "dark" in the description "A dark dinosaur moved from right to left." In contrast, under the same contextual conditions the adjective "large" was coded as context-unrelated when it was part of the description "A large gray dinosaur moved from my side to your side."
- An adjective which specified a value suggesting that the target object is around the center of the dimension was not coded as context-related. Consider for example an envelope target object which was paired with either a larger or a smaller envelope. The adjective "medium-sized" in the description "A medium-sized envelope moved from my side to your side" was coded as context-unrelated because it is not diagnostically related to one context object as opposed to the other.
- A simple mention of the dimension with no specific value did not satisfy the criteria for a context-related adjective. For example, the adjective "gray" in "A large gray dinosaur moved from my side to your side" was coded as context-unrelated because "gray" was not diagnostic with respect to the specific context object. The mention of the relative shade of gray was necessary in this set.

The mean adjectives count for the second presentation appears in Table 1. Based on our coding, we derived a dependent measure which reflects the use of context-related adjectives relative to the rest of the description. For each description, we calculated the relative use of context-related adjectives in the following way: the sum of the context-related adjectives was divided by the total number of words in the noun phrase identifying the critical

Table 1
Mean number of context-related and context-unrelated adjectives per description as a function of context condition and initiation speed (during second presentation)

	Initiation speed	
	Unspeeded	Speeded
Context-related		
Shared context	0.74	0.48
Privileged context	0.24	0.43
Context-unrelated		
Shared context	0.76	0.82
Privileged context	1.35	0.68

object, including all adjectives.[3] In this manner, we obtained a measure of the relative importance of the context-related information.

This measure is particularly useful when one considers first the general effect of speed pressure upon the descriptions. When speakers are put under pressure to quickly provide an utterance, they typically produce shorter utterances. Indeed, speeded subjects in our experiment gave shorter descriptions overall than subjects who were not speeded (means = 9.42 vs. 11.48 words per description, for the speeded and unspeeded conditions, respectively). This difference was significant, $F1(1, 20) = 4.68$, $p < .05$, $MS_e = 43.45$, $F2(1, 8) = 82.37$, $p < .001$, $MS_e = 76.39$ (analyses by subjects are indexed as *F1*, and analyses by items as *F2*). When descriptions shorten under pressure, adjectives are especially likely to be eliminated.[4] Indeed the mean number of adjectives per description dropped with speed (means = 1.49 vs. 1.10 adjectives per description for the unspeeded and speeded conditions, respectively). This drop was significant (though marginally with subjects), $F1(1, 20) = 3.72$, $p < .07$, $MS_e = 1.82$, $F2(1, 8) = 16.68$, $p < .004$, $MS_e = .188$.

Given that adjectives tend to drop under pressure, it might be more difficult for speakers to describe the objects in contrast to context because the adjective is the main tool to express such a contrast. If, in spite of this general tendency to drop adjectives, subjects use relatively *more* adjectives that express a privileged contrast under pressure, this constitutes strong evidence for the Monitoring and Adjustment model. Thus the relative use of the context-related adjectives, as described above, was chosen for analysis. These ratios were summed for each item in each cell and separately for each subject in each cell, and each item set was submitted to a 2 × 2 × 2 ANOVA involving three within-item factors: context (privileged vs. shared), initiation speed (unspeeded vs. speeded), and presentation (first vs. sec-

[3] The denominator of this ratio included the nouns as well because some utterances did not include adjectives; thus our ratio measure always had a denominator larger than zero.

[4] We are grateful to Kathryn Bock for pointing this out to us.

ond). The three-way interaction was significant, $F1(1, 20) = 6.96$, $p < .02$, $MS_e = .004$, $F2(1, 8) = 11.52$, $p < .01$, $MS_e = .003$. Given the significant three-way interaction, and that we are mainly interested in the descriptions for the second presentation, we focus on the analysis of the Context × Speed interaction for second descriptions.

Fig. 5 presents the extent to which speakers relied on context as indicated by the relative use of context-related adjectives. As predicted by the Monitoring and Adjustment model, while unspeeded speakers relied less on privileged than on shared context, speeded speakers relied on privileged and shared context to the same degree. When not speeded, the relative use of context-related adjectives was .29 and .09 for the shared and privileged contexts, respectively. This difference could be explained by both the Initial Design and the Monitoring and Adjustment models. The Initial Design model would suggest that speakers designed their utterances from the outset to incorporate common ground and consequently they used context more when it was shared than when it was privileged. In contrast, the Monitoring and Adjustment model explains this pattern as the result of the monitoring function which detected and corrected violations of common ground. The results of the speeded condition, however, supports the Monitoring and Adjustment model's interpretation and is inconsistent with that of the Initial Design model. The relative use of context-related adjectives was virtually identical with shared and privileged contexts when speakers were under time pressure to initiate their descriptions: the relative use of context-related

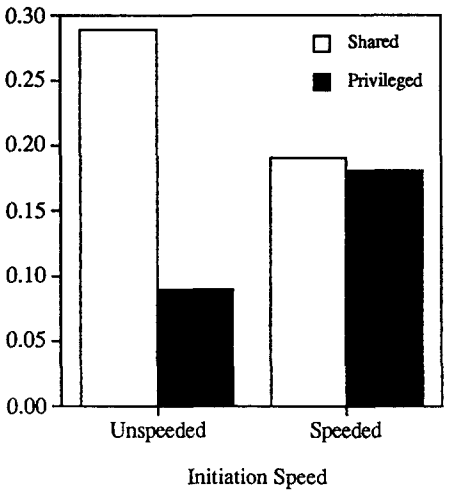

Fig. 5. Mean ratio of context-related adjectives to the total number of adjectives plus nouns per description as a function of context information and initiation speed, for the second presentation.

adjectives was .19 and .18 for the shared and privileged contexts, respectively. This interaction of context and speed was significant,[5] $F1(1, 20) = 3.6$, $p < .05$, $MS_e = .015$, $F2(1, 8) = 16.98$, $p < .01$, $MS_e = .005$. Most importantly, simple effects revealed that the two context conditions differed significantly with no time pressure, $F1(1, 20) = 8.09$, $p < .01$, $MS_e = .015$, $F2(1, 8) = 65.12$, $p < .001$, $MS_e = .003$, but they did not differ when speakers were speeded ($Fs > 1$).

We interpret the difference in the nature of descriptions between shared and privileged context, when unspeeded, to reflect the work of monitoring for violations of common ground. As converging evidence, we also examined the speakers' latency in beginning the descriptive noun phrases. In general, one should expect the mean latency to be shorter when subjects are speeded than when they are not. The Initial Design model predicts that the difference in latency between speeded and unspeeded subjects should be the same regardless of context type. It should not matter whether the context is privileged or shared. In contrast, the Monitoring and Adjustment model predicts a difference between the two context conditions. Overall, the speeded subjects should be faster than the unspeeded subjects, but this difference should be *larger* with privileged than with shared context. To understand this prediction, consider the way the model conceptualizes the different speed conditions. One could consider the speeded condition as a baseline because it produces descriptions which involve little or no monitoring as compared with the unspeeded condition when monitoring is allowed to take effect. Given that more revision work would be required with privileged context, the difference between the speed conditions should be larger with privileged than with shared context. In contrast to the Initial Design model which predicts only a main effect for speed, the Monitoring and Adjustment model also predicts a directional interaction between speed and context.

To test these predictions, we extracted from our raw data a measure of the onset latency for each of the relevant descriptions. In order to precisely measure the latency (with millisecond resolution), we converted the taped descriptions from analog to digital format by directly patching in the output from the tape player into a computer via an external 16-bit sound card. Using sound editing software at a 11.127 kHz sampling rate, we converted each description into a digital sound file, and each file was auditorily and spectrographically analyzed for two pieces of information: (1) the mouse-click indicating when the subject pressed the button to present the stimuli; and (2) the onset of the word that began the critical noun phrase, which was

[5] Given that the model predicts a *directional* interaction, and that the opposite pattern is not motivated by any theory, we used a directional test here and in the analysis of the reaction time data below. All other analyses, including simple effects, are nondirectional.

most often the first adjective.[6] We reasoned that the time difference between these two points represents the maximum latency the speaker took to formulate a description for the critical object. In this way, we obtained a measure of the time it took for subjects to formulate the descriptive noun phrase up to the point of initiation of its articulation; any additional processing necessary to take into account the listener's knowledge should be reflected in longer latencies.

The mean onset latencies for the second descriptions for each experimental group are presented in Table 2. The pattern of results is precisely what the Monitoring and Adjustment model predicts: while unspeeded subjects took longer than speeded subjects, the difference was more than twice as large with a privileged than with shared context (mean difference: 1597 ms and 727 ms, respectively). These times were submitted to a 2×2 (context: privileged vs. shared and initiation speed: unspeeded vs. speeded) ANOVA. The main effect of speed was significant, $F1(1, 20) = 47.58$, $p < .001$, $MS_e = 81014252$, $F2\ (1, 8) = 569.32$, $p < .001$, $MS_e = 12152230$. This confirms that subjects under pressure were indeed following instructions to initiate their descriptions quickly. Crucially, the interaction between time pressure and context information was also significant, $F1(1, 20) = 6.66$, $p < .01$, $MS_e = 1135162$, $F2(1, 8) = 74.99$, $p < .001$, $MS_e = 1702742$.

One might suggest that the speeded condition latency data should not serve as a baseline to test the hypothesis because this presupposes the basic assumptions of the model. A stronger test of the model would be to compare the shared and privileged context conditions only when subjects are not speeded. Indeed, unspeeded subjects took longer to initiate the noun phrase with privileged than with shared contexts (2935 ms and 2392 ms, respectively), which is consistent with the view that privileged contexts require extra monitoring work. This difference was significant, simple effect $F1(1, 20) = 5.20$, $p < .04$, $MS_e = 170279$, $F2(1, 8) = 40.17$, $p < .001$, $MS_e = 32463$. In summary, the results of the timing measure converge on the same conclusion as the results of the adjective count. When subjects are not

Table 2
Mean onset latencies for the critical descriptions in milliseconds, as a function of context condition and initiation speed (during second presentation)

	Context	
	Shared	Privileged
Unspeeded	2392	2935
Speeded	1665	1338
Difference	727	1597

[6] We purposely avoided measuring only up to the utterance onset because it was often drawn out or extended while the subjects appeared to be considering what to say. For example, the article "a" often became more like "Uhhhh".

pressed, they engage in monitoring for common ground, and only then one can see a difference between the extent to which they rely on shared and privileged contexts.

7. Discussion

The main result of our study is that speakers' descriptions relied on privileged context information less than on shared context information, but when speeded they relied on shared and privileged context to the same degree. We interpret this pattern of results as initial evidence against a model that assumes that common ground is part of an initial utterance design. If that model were correct, then speeded initiation of utterances should not result in a systematic disregard for common ground. We suggest that the pattern of our results supports a model that assumes that utterance plan is agnostic with respect to common ground information: Under time pressure, speakers relied on privileged and shared context to the same degree because they fell back on an initial plan which did not take common ground into account.

According to the Monitoring and Adjustment model, while the initial plan does not take common ground into account, speakers do monitor and attempt to correct and revise utterances which violate common ground. In the context of this study, the model assumes that the deadline procedure to initiate the description reduced or eliminated the effectiveness of the monitoring function, but did not affect the initial utterance plan. This effect was revealed both by speaker's adjective choice and latency to initiate the noun phrase.

What is the reason for such a selective effect on the production process? We assume that two factors could have contributed to the effect on monitoring but not on the initial plan: differential time course and differential reliance on resources. It is reasonable to assume that the initial planning of an utterance occurs fairly quickly whereas monitoring takes longer to initiate and bring to completion. For example, much of monitoring, such as monitoring for lexical errors, occurs relatively late in the production process, evaluating intermediate products of the utterance up to and following actual articulation (e.g., Levelt, 1983). This difference in the time course of planning and monitoring suggests that the relatively slower process – the monitoring – would be more likely than the initial planning to be affected by the reduction of initiation time. In addition to the difference in time course, it is reasonable to assume that the monitoring process is more prone to interference than the planning process because it is more dependent on available resources. This explanation assumes that the application of higher-level knowledge is more resource-dependent. To the extent that this holds, the monitoring process which applies information

about the other's knowledge should be more affected by the reduction of available resources (assuming fixed capacity, Just & Carpenter, 1993).

Why would it make sense for the production system to operate according to the Monitoring and Adjustment model? Perhaps because common ground is relatively high-level knowledge. Such knowledge might be too costly for the system to incorporate routinely in every utterance plan. By making common ground part of a correction mechanism the system might save some resources. This would be true only if utterances do not routinely require adjustments with respect to common ground. It is reasonable to assume that in many cases the utterance plan need not be corrected. This would occur when the information which is available to the speaker happens to be shared with the addressee. In such cases, when the speaker creates the initial plan and uses all information which is available to her, she also happens to use information which is part of the common ground. Consequently, the plan only uses information which is shared and no revision would be required. A typical case in point is when the speaker and the listener have the same referential set, as is true of most studies on referential communication (e.g., Ford & Olson, 1975; Issacs & Clark, 1987; Olson, 1970). For instance, both the speaker and the addressee might be looking at a large ball and a smaller ball. Both balls are part of the common ground of the interlocutors, and by definition they are "available" to the speaker. When the speaker attempts to plan a referential description of the smaller ball, she might take the contrast set into account and make reference to the size by saying "the small ball". According to the Monitoring and Adjustment model, the initial plan uses the contrast set neither because it is co-present for the addressee, nor because it is part of the common ground, but instead because it is available to the speaker. The fact that the contrast is accessible to the addressee plays no role in the initial planning of the utterance. In this case, because the contrast happens to be part of the common ground, subsequent monitoring would not intercept the utterance. In general, in many cases the plan might need no revision even though it did not initially use common ground information.

With the Monitoring and Adjustment model, a new notion of audience design emerges. If the system works like the model suggests, then there is no audience design in utterance planning. Some utterances might involve post hoc corrections, and some might involve no active audience design at all. Yet on the whole, utterances might still be appropriate for the audience, and they might look like they were pre-designed for that audience. The only other piece of evidence in line with this suggestion was reported by Brown and Dell (1987). They found that speakers tend to mention an atypical instrument more than a typical one. Though this tendency could be guided by a communicative attempt to be informative (Grice, 1975), it is not. Their subjects were as likely to mention the atypical more than the typical instrument regardless of whether or not the addressee was informed.

Apparently, the use of typicality of instrument is part of utterance planning, but not part of the audience design.

Would our results generalize beyond the descriptions of small circles and gray dinosaurs? Dell and Brown (1991) speculate that while knowledge about the listener ("model of the listener") might not play a role with "generic" knowledge such as typicality of an instrument, it might play a more substantial role with specific knowledge. In our experiment, the knowledge might have been more on the specific than the generic side. In contrast to typicality which is a relatively enduring quality, the contrast context in which our target objects were embedded were quite specific and ad hoc. The very fact that the speakers referred to the same object as "small" and "large" in different trials attests to the non-enduring quality of the information. This might suggest that speakers' disregard for a model of the listener applies for the generic as well as the specific case.

One might speculate that the role of the "model of the listener" might be different if interlocutors were allowed to collaborate (e.g., Clark & Schaefer, 1987; Clark & Wilkes-Gibbs, 1986) and listeners could provide feedback. As the evidence stands, though, it seems that our conclusions would only be strengthened if feedback were to be provided. For example, Schober (1993) found that speakers participating in a spatial perspective-taking task tended to use *more* egocentric perspectives when conversing with a partner than when talking to an imaginary addressee. This suggests that situations that do not involve full-scale collaboration might cause speakers to be more careful in considering their listener's knowledge. Krauss and Fussell (1991) suggest precisely this: "speakers in interactive contexts may feel less need to consider the addressee's knowledge because they know that the listener can ask questions to clarify meanings as necessary" (p. 19). If this is true, then our experiment may have overestimated the extent to which speakers use common ground.

In evaluating the model that we propose one must be cautious not to overgeneralize to all aspects of common ground. Our study directly tested the model using a particular instantiation of common ground: physical co-presence (Clark and Marshall, 1981). We compared the case where the context objects were co-present for the interlocutors to the case where they were not co-present. Future research could explore the generalizability of the model to other instantiations of common ground, such as common information established by linguistic co-presence or by community membership.

The role of common ground by virtue of linguistic co-presence has been investigated in a comprehension context, and initial evidence suggests a strong similarity between the way common ground functions in production and the way it is used in comprehension. Keysar and Paek (1993) demonstrated that listeners do not incorporate common ground in their interpretations at the outset. When they understand definite references, they attempt

to identify a unique referent regardless of whether or not it is accessible to the speaker. When the interpretation used inaccessible referents, comprehension was delayed, probably because of the need to correct the interpretation. It seems that common ground comes into play only if the comprehension system fails to resolve the referential description.

Third-party observers' use of common ground appears to follow an analogous model as well. Keysar (1994) showed that when readers understand a speaker's intention, they tend to attribute to an addressee the perception of that intention. This happens even when the utterance is clearly ambiguous and even though they know that the addressee lacks the disambiguating information which they possess as readers. Keysar, Paek and Balin (under review) presents evidence which supports a model for this phenomenon which is analogous to the Monitoring and Adjustment model in production: when readers or observers attempt to evaluate the interpretations of uninformed addressees, they quickly interpret the utterance from their own perspective. In order to attribute a perception of an intention to the addressee, they make adjustments by taking into account which information is accessible and which is inaccessible to the addressee. As with monitoring in production, this adjustment process is more susceptible to external interference, which occasionally results in under-adjustment.

8. Conclusion

Our study allows us to provide an initial characterization of the role of common ground in utterance production. Audience design which uses common ground information might not be inherent to routine utterance planning. According to this view, some utterances could be well designed with minimal use of common ground. Many utterances which look like they are designed for their audiences are not so designed, but only happen to appear as such. Some utterances, whose plans violate common ground, are retroactively revised to accommodate the audience. This correction is part of the monitoring function of utterance production.

Acknowledgements

The study reported here was presented to the University of Chicago in partial fulfillment of William Horton's MA degree requirements. We are grateful for the financial support provided by PHS, grant R29 MH49685 to the University of Chicago, Boaz Keysar Principal Investigator, to the Social Sciences Division for an SSDR research grant, and to the Jacob K. Javits Fellowship program for a graduate fellowship supporting William Horton. We are indebted to Linda Ginzel, David McNeill and Howard Nusbaum for insightful comments on earlier drafts and to Carol Laurent and Erin Pitts for

technical help. Finally, we would like to thank three anonymous reviewers for their help and encouragement.

References

Ackerman, B.P., Szymanski, J., & Silver, D. (1990). Children's use of the common ground in interpreting ambiguous referential utterances. *Developmental Psychology, 26*, 234–245.

Bock, J.K. (1986). Syntactic persistence in language production. *Cognitive Psychology, 18*, 355–387.

Bock, K. (1990). Structure in language: creating form in talk. *American Psychologist, 45*, 1221–1236.

Bock, K. (1995). Sentence production: from mind to mouth. In J.L. Miller & P.D. Eimas (Eds.), *Handbook of perception and cognition: Vol. 11: Speech, language, and communication* (pp. 181–216). Orlando, FL: Academic Press.

Bock, J.K., & Kroch, A.S. (1989). The isolability of syntactic processing. In G.N. Carlson & M.K. Tanenhaus (Eds.), *Linguistic structure in language processing*. Dordrecht: Kluwer.

Bock, K., Loebell, H., & Morey, R. (1992). From conceptual roles to structural relations: bridging the syntactic cleft. *Psychological Review, 99*, 150–171.

Brown, P.M., & Dell, G.S. (1987). Adapting production to comprehension: the explicit mention of instruments. *Cognitive Psychology, 19*, 441–472.

Butterworth, B. (1989). Lexical access in speech production. In W.D. Marslen-Wilson (Ed.), *Lexical representation and process*. Cambridge, MA: MIT Press.

Butterworth, B., & Hadar, U. (1989). Gesture, speech, and computational stages: a reply to McNeil. *Psychological Review, 96*, 168–174.

Clark, H.H. (1992). *Arenas of language use*. Chicago: University of Chicago Press.

Clark, H.H., & Haviland, S.E. (1977). Comprehension and the given-new contract. In R.O. Freedle (Ed.), *Discourse production and comprehension*. Norwood, NJ: Ablex.

Clark, H.H., & Marshall, C.R. (1981). Definite reference and mutual knowledge. In A.H. Joshe, B. Webber, & I.A. Sag (Eds.), *Elements of discourse understanding*, Cambridge, UK: Cambridge University Press.

Clark, H.H., & Murphy, G.L. (1982). Audience design in meaning and reference. In J.F. LeNy & W. Kintsch (Eds.), *Language and Comprehension*. Amsterdam: North-Holland.

Clark, H.H., & Schaefer, E.F. (1987). Concealing one's meaning from overhearers. *Journal of Memory and Language, 26*, 209–225.

Clark, H.H., Schreuder, R., & Buttrick, S. (1983). Common ground and the understanding of demonstrative reference. *Journal of Verbal Learning and Verbal Behavior, 22*, 245–258.

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

Cooper, W.E., & Paccia-Cooper, J. (1980). *Syntax and Speech*. Cambridge, MA: Harvard University Press.

Cutler, A. (1980). Errors of stress and intonation. In V.A. Fromkin (Ed.), *Errors in linguistic performance: Slips of the tongue, ear, pen, and hand*. New York: Academic Press.

Cutler, A., & Isard, S. (1980). The production of prosody. In B. Butterworth (Ed.), *Language production: Vol. 1. Speech and talk*. London: Academic Press.

Dell, G.S. (1986). A spreading activation theory of retrieval in sentence production. *Psychological Review, 93*, 283–321.

Dell, G.S., & Brown, P.M. (1991). Mechanisms for listener-adaptation in language production: limiting the role of the "model of the listener". In D. Napoli & J. Kegl (Eds.), *Bridges between psychology and linguistics*. New York: Academic Press.

Dell, G.S., & O'Seaghdha, P.G. (1991). Mediated and convergent lexical priming in language production: a comment on Levelt et al. (1991). *Psychological Review, 98*, 604–614.

Dell, G.S., & O'Seaghdha, P.G. (1992). Stages of lexical access in language production. *Cognition, 42*, 287–314.

De Smedt, K., & Kempen, G. (1987). Incremental sentence production, self correction, and coordination. In G. Kempen (Ed.), *Natural language generation: Recent advances in artificial intelligence, psychology and linguistics*. Dordrecht: Kluwer.

Deutsch, W., & Pechmann, T. (1982). Social interaction and the development of definite descriptions. *Cognition, 11*, 159–184.

Ferreira, F. (1993). The creation of prosody during sentence production. *Psychological Review, 100*, 233–253.

Ford, W., & Olson, D.R. (1975). The elaboration of the noun phrase in children's description of objects. *Journal of Experimental Child Psychology, 19*, 371–382.

Fromkin, V. (1971). The nonanomalous nature of anomalous utterances. *Language, 47*, 27–52.

Garrett, M.F. (1976). Syntactic processes in sentence production. In R.J. Wales & E. Walker (Eds.), *New approaches to language mechanisms*. Amsterdam: North-Holland.

Garrett, M.F. (1988). Processes in language production. In F.J. Newmeyer (Ed.), *Linguistics: The Cambridge survey, Vol. 3. Language: Psychological and biological aspects*. Cambridge, UK: Cambridge University Press.

Glucksberg, S., Krauss, R.M., & Higgins, E.T. (1975). The development of referential communication skills. In F.E. Horowitz (Ed.), *Review of child development research*. Chicago: University of Chicago Press.

Glucksberg, S., Krauss, R.M., & Weisberg, R. (1966). Referential communication in nursery school children: method and some preliminary findings. *Journal of Experimental Child Psychology, 3*, 333–342.

Grice, H.P. (1975). Logic and conversation. In P. Cole & J. Morgan (Eds.), *Syntax and semantics 3: Speech acts*. New York: Academic Press.

Harley, T.A. (1993). Phonological activation of semantic competitors during lexical access in speech production. *Language and Cognitive Processes, 8*, 291–309.

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

Just, M.A., & Carpenter, P.A. (1993). A capacity theory of comprehension: individual differences in working memory. *Psychological Review, 99*, 122–149.

Keysar, B. (1994). The illusory transparency of intention: linguistic perspective taking in text. *Cognitive Psychology, 26*, 165–208.

Keysar, B., & Paek, T.S. (1993). Definite reference and mutual ignorance. In *The 34th annual meetng of the Psychonomics Society*, Washington, DC, November 1993.

Keysar, B., Paek, T.S., & Balin, J.A. (under review). From principles to processes in theories of language use: A perspective adjustment model.

Krauss, R. (1987). The role of the listener: addressee influences on message formulation. *Journal of Language and Social Psychology, 6*, 81–98.

Krauss, R.M., & Fussell, S.R. (1991). Perspective-taking in communication: representations of others' knowledge in reference. *Social Cognition, 9* 2–24.

Krauss, R.M. & Glucksberg, S. (1977). Social and nonsocial speech. *Scientific American, 236*, 100–105.

Laver, J.D.M. (1980). Monitoring systems in the neurolinguistic control of speech production. In V.A. Fromkin (Ed.), *Errors in linguistic performance*. New York: Academic Press.

Levelt, W.J.M. (1983). Monitoring and self-repair in speech. *Cognition, 14*, 41–104.

Levelt, W.J.M. (1989). *Speaking: From intention to articulation*. Cambridge, MA: MIT Press.

McNeill, D. (1989). A straight path – to where? Reply to Butterworth and Hadar. *Psychological Review, 96*, 175–179.

Meyer, A.S. (1994). Timing in sentence production. *Journal of Memory and Language, 33*, 471–492.

Olson, D.R. (1970). Language and thought: aspects of a cognitive theory of semantics. *Psychological Review, 77*, 257–273.

Schober, M.F. (1993). Spatial perspective-taking in conversation. *Cognition, 47*, 1–24.

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

Schriefers, H., Meyer, A.S., & Levelt, W.J.M. (1990). Exploring the time course of lexical access in language production: picture–word interference studies. *Journal of Memory and Language*, *29*, 86–102.