Definite Reference and Mutual Knowledge: Process Models of Common Ground in Comprehension

Boaz Keysar, Dale J. Barr, and Jennifer A. Balin

The University of Chicago

and

Timothy S. Paek

Stanford University

What role does mutual knowledge play in the comprehension process? We compare two answers to this question for the comprehension of definite reference. The *Restricted Search hypothesis* assumes that addressees rely on the principle of optimal design and understand definite reference by restricting the search for referents to entities in common ground. The *Unrestricted Search hypothesis* assumes that the search for referents is not restricted to entities in common ground. Only the Unrestricted Search hypothesis predicts that entities that are not in common ground would interfere with comprehension of definite reference. Experiment 1 reveals such interference in increased errors and verification latencies during the resolution of pronouns. Experiment 2 demonstrates the interference by tracking the addressee's eye movements during the comprehension of demonstrative reference. We discuss alternative models of comprehension that could account for the results, and we describe the role that common ground plays in each model. We propose a Perspective Adjustment model that assumes a search for referents that is independent of common ground, coupled with a monitoring process that detects violations of common ground and adjusts the interpretation. This model assumes a role for common ground only when a correction is needed. We challenge both the assumption that addressees follow the principle of optimal design and the assumption that the principle is optimal.   © 1998 Academic Press

Key Words: language; communication; perspective; interpretations; definite reference; process models.

''How about some pepper?'' is a pragmatically ambiguous question. The speaker may be proposing to spice up an article, asking whether or not to add pepper to a grocery list, or making an opaque response to a Speech-Act theorist who uses the infamous example, ''can you pass the salt?'' In the same way, *any* utterance is ambiguous because the speaker could be using it to convey a variety of possible intentions. This imposes a potentially heavy burden on participants in conversation (i.e., interlocutors) because in order to communicate effectively they need to minimize ambiguity.

To explain how interlocutors disambiguate intentions, pragmatic theories outline a variety of principles. Grice (1975; Sadock, 1978) suggests that interlocutors assume mutual adherence to the *cooperative principle*. This assumption reduces ambiguity for listeners and consequently simplifies their task. Sperber and Wilson (1986, 1987) argue that the single principle of *relevance* is sufficient to account

for the way expressions are interpreted. Many scholars argue that interlocutors use their *common ground* to communicate effectively— speakers use shared information in constructing their utterances, and addressees use it to disambiguate speakers' intentions (Clark & Carlson 1981; Clark & Marshall, 1981; Clark, Schreuder, & Buttrick, 1983; Fussell & Krauss, 1989; Gerrig & Littman, 1990; Gibbs, Mueller, & Cox, 1988; Krauss & Fussell, 1991, in press). For example, Fussell and Krauss (1989) demonstrated that speakers tailor messages for friends and Clark, Schreuder, and Buttrick (1983) argue that addressees use mutually salient information to disambiguate a referring expression. Though the initial proposal of the role of common ground by Clark and Carlson (1981, 1982) was strongly challenged (e.g., Johnson-Laird, 1982; Sperber, 1982; Sperber & Wilson, 1982), it is now generally accepted that common ground plays an important role in comprehension.

Though the cognitive study of mental representations and mental processes go hand in hand, the literature on pragmatics in general and the study of common ground in particular do not follow this tradition. Some papers focus on the nature of the representation of mutual knowledge (e.g., Clark & Carlson, 1981; Clark & Marshall, 1981). They describe a variety of sources for common ground and outline a possible memory representation in the form of ''diary entries'' and ''model of the other'' which spells out what information is mutually known with that other. Other studies investigate the role the represented information plays in comprehension and production (e.g., Clark, Schreuder, & Buttrick, 1983; Fussell & Krauss, 1989; Gerrig & Littman, 1990). Further studies consider the interpersonal processes which are involved in conversation, such as the way people collaborate and coordinate their beliefs when they converse (e.g., Clark & Schaefer, 1989; Wilkes-Gibbs & Clark, 1992). There are assumptions in the field about the way common ground affects processing, but these assumptions have not been tested. Consequently, there has not been a critical evaluation of alternative process models of the role of common ground in comprehension. While we know that listeners represent common ground information, and we know that they use that information, we do not know *how*. Our goal is to evaluate the role that common ground plays in comprehension.

THE CONCEPT OF COMMON GROUND

Common ground is defined as a type of shared information: ''The common ground between Ann and Bob, for example, is the sum of their mutual knowledge, mutual beliefs, and mutual suppositions'' (Clark, 1992, p. 3). A piece of information p is mutually believed by Ann and Bob if each one of them believes p and each one believes that the other believes it and so on. The notion of mutuality is a central element in the definition: It is not sufficient for information to be independently known to each interlocutor. This by itself will not make it part of their common ground. It must also be mutually known by them, or at least they should have a very good reason to believe that it is mutually known (e.g., Clark & Marshall, 1981). Mutual knowledge, then, is a type of meta-knowledge; it is knowledge about knowledge. The question is, given that a comprehender knows that information is mutually known with a speaker, what role does this meta-knowledge play in the process of understanding that speaker?

The general assumption in the field is that such mutual knowledge or common ground plays an important role in virtually any act of comprehension. Clark and Carlson (1981), for example, argue that common ground is used in the comprehension of conventional expressions, speech acts, definite reference, and contextual expressions. With respect to common ground, definite reference has been by far the most extensively studied. We will therefore focus on processing assumptions regarding the role that common ground plays in the comprehension of definite reference.

DEFINITE REFERENCE AND COMMON GROUND

People often use definite descriptions to establish reference. It can be accomplished with

different linguistic forms: the definite article, anaphora, demonstrative reference, and so on. Although definite description has a variety of functions in discourse, we will focus on its role in identifying unique individuals. Definite descriptions do not always refer to specific individuals, but are sometimes ''attributive'' (e.g., Donnellan, 1966; Johnson-Laird & Garnham, 1980; Mueller-Lust & Gibbs, 1991; Ortony & Anderson, 1977). For example, ''The world's inhabitants at the year 3000 will have no toes'' includes an attributive definite description which means ''whoever those people may be.'' In contrast, a referential description, as in ''today's inhabitants of the world have toes,'' picks out specific individuals. Given that we are interested in the reference-picking property of definite descriptions, we do not deal with attributive descriptions. Similarly, we are not concerned with the generic use of definite reference, as in ''the monkey is not a bilingual animal,'' where ''the monkey'' does not refer to a particular monkey. Instead we focus on and use specific definite reference as in ''this monkey has no idea what she is talking about.''

In a classic paper, Chafe (1976, p. 38) argued that the central use of a definite description suggests *identifiability*—that the listener is able to uniquely identify a referent. Thus, when speakers use a definite reference, listeners identify the intended referent by searching for a unique referent. But how do they do that and what role does common ground play in the process of identifying the intended referent? Different from the traditional approach in the study of language use, our ultimate goal is to specify a process model of the way intentions are communicated. In this paper, we contrast two possible hypotheses regarding the role that common ground might play in such a processing model.

THE ROLE OF COMMON GROUND IN THE COMPREHENSION PROCESS

When Boris' wife says to him in the morning, ''he is awake now,'' Boris identifies the referent of the pronoun ''he'' as their son David. Boris uses two pieces of information for this interpretation: He uses the semantic knowledge that the pronoun ''he'' can pick out any male person, among them his son, and he uses the pragmatic information that David is in common ground with his wife. In order to identify the intended referent, Boris searches his memory for potential referents. The question is, does the higher level pragmatic knowledge of common ground constrain that search to entities which are in common ground?

Whether or not higher level knowledge constrains lower level processing is an issue which has been debated across the board in cognitive psychology in general and psycholinguistics in particular. For example, in the study of word recognition, there is a debate over the extent to which different types of contexts can constrain lexical access (e.g., Tanenhaus & Lucas, 1987). Some models suggest that sentential context does not constrain lexical access (e.g., Hogaboam & Perfetti, 1975; Kintsch & Mross, 1985; Onifer & Swinney, 1981; Swinney, 1979). Other models claim that context can constrain lexical access to the contextually appropriate meaning (e.g., Glucksberg, Kreuz, & Rho, 1986; Simpson, 1981; Tabossi, Colombo, & Job, 1987).

By analogy to the difference between lexical access models, we will contrast two hypotheses regarding the comprehension of definite reference. One hypothesis assumes that the search for referents is constrained by common-ground knowledge to referents in common ground. In contrast, the other hypothesis assumes no such restricted search.

THE ''RESTRICTED SEARCH'' HYPOTHESIS

One possibility is that the search for referents is restricted to entities that are in common ground. So, when a friend says to you ''the dog,'' you will be searching for referents among dogs which are in your common ground with that friend. The search will not include the stray dog which you saw right before you met your friend, assuming no evidence that the dog is in your common ground. This hypothesis requires that common-ground information would be identified as such in memory. Indeed, this has been a common assumption in the field. For example, Clark

and Marshall (1981; also Greene, Gerrig, McKoon, & Ratcliff, 1994) invoke the metaphor of reference diaries to describe how common ground is referenced in memory to comprise models of the other: ''Ann's model of Bob, in short, contains just those parts of her diary and encyclopedia that will be useful for getting him to understand her. . .It will also contain just those parts that will allow her to understand him.'' (Clark & Marshall, 1981, p. 55).

It makes sense that listeners restrict their search to common-ground referents because speakers are supposed to follow the principle of optimal design (Clark, Schreuder, & Buttrick, 1983): ''The speaker designs his utterance in such a way that he has good reason to believe that the addressee can readily and uniquely compute what he meant on the basis of the utterance along with the rest of their common ground.'' (p. 246). If listeners assume that speakers adhere to this principle, then they might search only among common ground referents.

Clark and Carlson (1981) argue that such a restricted search is preferable: ''When a listener tries to understand what a speaker means on some occasion, it would be advantageous if the process he uses could limit what it retrieves from memory to some portion of the total information that could be made available. In particular, it should limit itself to the intrinsic context, that portion of the information that may be needed for the process to succeed.'' (p. 319). According to Clark and Carlson, then, common ground becomes the context for comprehension: ''Our proposal is straightforward: *The intrinsic context for understanding what a speaker means on some occasion is the common ground that the listener believes holds at that moment between the speaker and the listeners he or she is speaking to.*'' (p. 319; emphasis in original).

The question we started with is: What role does knowing who knows what play in comprehension? The answer that the Restricted Search hypothesis provides is straightforward: The role of mutual knowledge is to restrict the search for referents to entities in common ground. According to this hypothesis, then, pragmatic knowledge of common ground guides the search for referents from the outset.

THE ''UNRESTRICTED SEARCH'' HYPOTHESIS

In contrast to the Restricted Search hypothesis, this hypothesis assumes that when addressees understand definite reference they conduct a search for referents which is not guided by mutual knowledge. When the addressee understands ''I saw Mary last night,'' the unrestricted search will pick out as a potential referent an available ''Mary'' whether or not that particular Mary is mutually known to the speaker and the addressee. So, the answer that this hypothesis provides for our question is: The meta-knowledge that an entity is part of common ground does not play a guiding role in the search for referents of definite reference. Therefore, this hypothesis assumes a search for referents which can be characterized by any theory of referential understanding which does not make assumptions about common ground (e.g., Gernsbacher, 1989; McKoon & Ratcliff, 1992; and others).

This hypothesis might seem strange given a vast literature that assumes that language users follow the principle of optimal design and thus rely on mutual knowledge in the understanding of definite reference. Yet the Unrestricted Search hypothesis is motivated by recent findings from our laboratory that demonstrate that under certain conditions language users violate the principle of optimal design (e.g., Horton & Keysar, 1996; Keysar, 1994; Keysar, in press; Keysar, Barr, & Horton, in press).

Given our previous findings, we have proposed that language users do not rely on common ground unless they make an error. For example, Horton and Keysar (1996) proposed a ''Monitoring and Adjustment'' model for production (see also Dell & Brown, 1991). They demonstrated that when speakers plan utterances, they do not rely on their mutual knowledge with their addressee; instead, they plan their utterances with no regard to such meta-knowledge. Mutual knowledge does

play a role in the monitoring process; speakers monitor their utterance plans and if they detect a violation of common ground they revise those plans. In the general discussion we will propose an analogous model for comprehension—the Perspective Adjustment model.

THE CRITICAL TEST OF THE HYPOTHESES

The experiments are designed to compare the two hypotheses about the role of common ground in the comprehension of definite reference. The two hypotheses differ with respect to the nature of the search for referents. The critical difference is that the Restricted Search hypothesis assumes that only referents in common ground are being considered, whereas the Unrestricted Search hypothesis assumes that the search is not restricted to entities in common ground. The critical test case, then, is whether referents which are *not* in common ground are included in the search and occasionally picked as referents. The Restricted Search hypothesis predicts that they would not be included in the search, while the Unrestricted Search hypothesis predicts that they would.

Experiments 1 and 2 perform this critical test with participants who are actual addressees. The basic logic is the following. Both experiments differentiate between (1) information that is in common ground, *and* (2) information known only to the addressee and thus clearly not in common ground. This distinction observes the Subsuming Theory criterion (Keysar, 1997) and allows us to investigate whether addressees understand a definite reference via a search which is or is not restricted by common ground.

In contrast to the Restricted Search hypothesis, then, the Unrestricted Search hypothesis predicts a systematic error pattern in the understanding of definite reference. Clark and Carlson (1981) allude to such differential predictions when they propose a restricted search and discuss its alternative:

> What we have proposed is that when a listener tries to understand what a speaker means, the process he goes through can limit memory access to information that is common ground between the speaker and his addressees. At the very least, it must distinguish between information that is and is not part of the common ground, because otherwise in certain situations it will systematically misinterpret conventions, direct and indirect speech acts, definite reference, and contextual expressions. (p. 328)

Our experiments test precisely for such systematic misinterpretations. The Restricted Search hypothesis does not predict such systematic errors, but the Unrestricted Search hypothesis does.

EXPERIMENT 1: WHO IS THE REFERENT OF THE PRONOUN?

Many studies have investigated the way addressees interpret anaphoric reference (e.g., Caramazza, Grober, Garvey, & Yates, 1977; Dell, McKoon, & Ratcliff, 1983; Ehrlich, 1980; Rinck & Bower, 1995). There is a general consensus that the availability of a referent is important for its quick identification as the referent of an anaphoric reference, such as a pronoun (e.g., Gernsbacher, 1989, 1990; McKoon, Ward, & Ratcliff, 1993) or an anaphoric verb phrase (Malt, 1985; Murphy, 1985). Yet none of the available studies shed light on the question in focus because they did not involve a diverging perspective for the participants and the ''speakers.'' Experiment 1, using pronoun resolution, tests the different predictions of the two hypotheses.

How can the resolution of a pronominal reference provide the test between the hypotheses? Consider the following situation. It is evening, and Boris' young daughter is playing in the other room. Boris, who lives in Chicago, is thinking of calling his lover in Europe. He decides not to call because she is probably asleep given the transatlantic time difference. At that moment his wife returns home and asks, ''Is she asleep?'' How would Boris search for the referent of ''she''? The Restricted Search hypothesis assumes that the search for the referent of ''she'' is limited to entities which are in common ground. Consequently, it predicts that the lover would not be a potential referent for the pronoun because Boris' wife is not informed about the lover.

In contrast, the Unrestricted Search hypothesis assumes that the search for referents is not restricted to common ground entities. Instead, Boris searches for available referents for the pronoun ''she.'' Given that the sleeping lover was readily available, the Unrestricted Search hypothesis predicts that Boris would quickly interpret ''she'' to refer to her. Experiment 1 is modeled after this ''real-life'' analog with pronoun interpretation.

Participants played a question and answer game with another participant. They received the information that John read a newspaper, and then their interlocutor asked them, ''Did he read a novel?'' The critical issue was: How long would it take them to answer the question? In addition, they performed a secondary memory task. They received information which was clearly labeled as inaccessible to the speaker and whose only goal was supposed ''to make the task a bit more difficult by increasing memory load.'' In this case this information was either ''Ralph read a novel'' or ''Mary read a novel.'' If participants comprehended ''Did he read a novel?'' without being restricted to common ground, then they should consider ''Ralph'' as a potential referent to the pronoun ''he.'' This predicts interference when the privileged (i.e., memory task) sentence is about Ralph compared with the case when it is about Mary. To complete the analogy, when Boris' wife asked, ''Is she asleep?'' the Unrestricted Search hypothesis predicts interference only if Boris' secret lover is female, not if the secret lover is male. In contrast, if the search for referents is restricted to common ground entities as the Restricted Search hypothesis assumes, then entities which are definitely not in common ground (e.g., Ralph) would not be considered.

Method

Participants. Forty-eight native English speakers played the role of addressees in the experiment for pay. None had participated in a similar experiment before. One participant was replaced because of experimenter error and five participants were replaced because they made errors that exceeded criterion (i.e.,

more than four errors or more than four recall errors), suggesting that they were not following instructions.

Experimental setup and procedure. The experimenter explained that participants would play an information game with another participant. The ''other participant'' was a confederate and she always played the role of the ''speaker,'' while the real participant always played the role of the ''addressee.'' The experimenter then conducted a lengthy preexperiment session with both speaker and addressee present to ensure that participants believed that the confederate was also a real participant. In order to make sure they understood the game, participants received instructions and practice and got to play the role of the speaker as part of the preexperiment training. The postexperiment interview verified that there was no suspicion about the ''other participant,'' without exception. Addressees were told that their goal in the game was to use the information provided by the experimenter to answer questions posed by the ''speaker.'' During the actual experiment, participants–addressees saw this information on the screen and heard prerecorded questions by the confederate–speaker. They all believed that they were interacting with the other participant via an intercom.

To motivate the questions, the experimenter explained that the speaker had a scenario which ended with a question and provided the following example:

> Joe and Rachel are the only two workers for a small delivery company in downtown Chicago. Joe makes speedy deliveries by bicycle, which allows him to avoid heavy traffic. Rachel drives the company truck. One Monday, the company made two deliveries: a sofa and a cake. Rachel delivered something. What did she deliver?

Participants were told that when the ''speaker'' receives the scenario, the ''addressee'' will receive further information about Rachel, in this case: ''Rachel delivered the sofa.'' This information was always the answer to the question that ended the speaker's scenario. The speaker's goal was to find

out the answer to that question. So in this case, the speaker asked, ''Did Rachel deliver the sofa?'' and the addressee was supposed to answer ''yes.''

In addition to the main task of the experiment—answering the partner's question—the experimenter explained that because the task is too easy, they would receive an additional sentence that they would have to remember. For example, with the delivery scenario the second sentence was ''Marla delivered a cake.'' The second sentence was framed as an additional load on their memory. To make sure that participants did not misunderstand, the instructions stressed several times that the second sentence was not relevant to the main task because its protagonist (e.g., Marla) does not appear in the scenario, and therefore the second sentence should not be used to answer the partner's question. (See Appendix A for instructions.)

To ensure that participants were not confused about the accessibility status of the information (i.e., about who knows what), the first sentence was always the scenario-relevant one and the second one was always the memory-load sentence. In addition, participants received the first sentence on the screen and were encouraged to take as much time as they wanted to memorize it. When they were ready, they pressed a ''continue'' button to replace the first sentence with the second sentence. Again, they took as much time as they wanted to memorize it. When they were ready, they pressed the ''continue'' button to hear their partner's question. They answered ''yes'' or ''no'' by pressing the corresponding key. The instructions stressed both speed and accuracy. To make sure participants kept both sentences in mind, they were instructed to recall one of them at random after the end of each trial. They recalled the sentence verbally into a microphone.

Materials. Experimental items included two sentences, appearing consecutively on the screen, and all items required a ''no'' response because of the information provided by the first sentence. For example, when ''John was bitten by a cobra'' was the first sentence, the speaker's question was ''Was he bitten by a rat?'' and so the correct answer was ''no.'' The first sentence in each item set was fixed. The second sentence included information about a character who played no role in the (presumed) speaker's scenario. In this example, the second sentence was ''Helen was bitten by a rat.'' In addition, the question used either a name or a pronoun, e.g., ''Was John bitten by a cobra?'' or ''Was he bitten by a cobra?'' The experiment included 16 different item sets, all modeled after this example. Half the items used a male name and the other half used a female name in the first sentence. The items were divided into four different blocks, each block included one version of each item set, a fourth in each condition. Actor's gender and question format were counterbalanced across blocks. Each participant received only one block. A similar number of fillers were used in order to prevent response strategies.

Sixteen fillers required a ''yes'' response because the first sentence provided a positive answer to the question. For example, the first sentence of a filler was ''Dillon was late for a discussion,'' with the corresponding question ''Was Dillon late for a discussion?'' These fillers were modeled after experimental items and were the same in all blocks. Half had a male actor and half a female actor in the first sentence. In addition, half of the fillers had a question that used a name and the rest had a question that used the corresponding pronoun. Finally, half the fillers had a same-gender actor for the second sentence and the other half had a different-gender actor. Because the fillers were modeled after items in this way, participants could not have known in advance whether the answer was ''yes'' or ''no.'' To preempt possible response strategies by participants, the blocks included an additional set of sixteen fillers, half negative and half positive. For example, to make sure participants paid attention to the verb in the question, some of these fillers used a different verb than the one in the first sentence. Additionally, some of the fillers included a second sentence that could have caused interference.

Design. We manipulated two factors: (1)

The actor of the privileged sentence either was of the same gender or a different gender than the actor in the speaker's question (e.g., with ''he'' in the question, the privileged sentence had either Rob or Helen as the actor). (2) The question used either the name of the actor of the first sentence or the corresponding personal pronoun (e.g., either ''John'' or ''he''). This yielded a 2 (Gender: Same or Different) × 2 (Question Format: Name or Pronoun) within-subjects design.

Predictions. The Unrestricted Search hypothesis makes a clear prediction regarding the Gender condition when the question includes a pronoun. An unrestricted search would consider referents from the privileged sentence. So when the question uses a pronoun, ''Was he bitten by a rat?'' addressees should have a tendency to respond ''yes'' when the privileged sentence has a same-gender actor (''Rob was bitten by a rat'') than when it has a different-gender actor (''Helen was bitten by a rat''). This should cause interference, delay a correct ''no'' response and increase the error rate. In contrast, the Restricted Search hypothesis predicts no difference between the two conditions because the search will not include entities which are not in common ground.

Results and Discussion

Analysis of response latency. We measured latency from the beginning of the recorded message to the response of the participant. The data for correct responses did not include reaction times which were longer than 4.5 *SD* above the mean, which eliminated less than 1% of the data. The pattern of the means for the four cells supports the Unrestricted Search hypothesis. (See Fig. 1). When the question used a pronoun, addressees were on average 170 ms slower to respond when the privileged sentence included a same-gender actor than a different-gender actor. In contrast, same- and different-gender conditions did not differ when the question used a proper name (Means = 2016 and 2033 ms, respectively).

The reaction time data were averaged for each participant in each cell, and for each item

in each cell and were submitted separately to a 2 (Gender: Same or Different) × 2 (Question Format: Name or Pronoun) ANOVA with repeated measures. For both measures we report analyses over participants as *F1* and over items as *F2*. The use of name vs. pronoun had no significant effect, $F1(1,45) = 1.30$, $p = .3$, $MSE = 97642$, $F2 < 1$. Because the overall mean of same-gender sentence was higher than different-gender sentence, Gender had a significant effect, $F1(1,45) = 4.97$, $p < .05$, $MSE = 83212$, $F2(1,15) = 5.01$, $p < .05$, $MSE = 18503$. Because Gender had an effect only in the pronoun condition, the interaction was significant, $F1(1,45) = 4.25$, $p < .05$, $MSE = 80667$, $F2(1,15) = 6.69$, $p < .05$, $MSE = 22168$. The slowdown when the privileged sentence had an actor which could be the referent of the pronoun suggests that addressees consider that actor as a referent and reject it. This pattern is precisely what the Unrestricted Search hypothesis predicts for addressees. The Restricted Search hypothesis cannot account for these results.

Analysis of error rates. The second dependent measure of the predicted interference in comprehension was the extent to which addressees made errors and responded in the affirmative instead of the negative. Participants' overall accuracy was relatively high (93% accurate) but the pattern of means or errors per cell again supports the Unrestricted Search hypothesis. (See Fig. 2). Mirroring the pattern of reaction-time data, participants made an average of 10 percentage points more errors when a pronoun question was coupled with a privileged sentence that used a same-gender actor rather than a different-gender actor. Questions that used a proper name did not involve a different error rate for same- vs different-gender actor (Means = 5% for both).

The error-rate data were submitted to a 2 (Gender: Same or Different) × 2 (Question Format: Name or Pronoun) ANOVA with repeated measures. Question Format had no significant effect with participants as a random factor, $F1(1,47) = 1.78$, $p = .18$, $MSE = .0262$, but was significant over items, $F2(1,15) = 4.66$, $p < .05$, $MSE = .0034$. Gender had
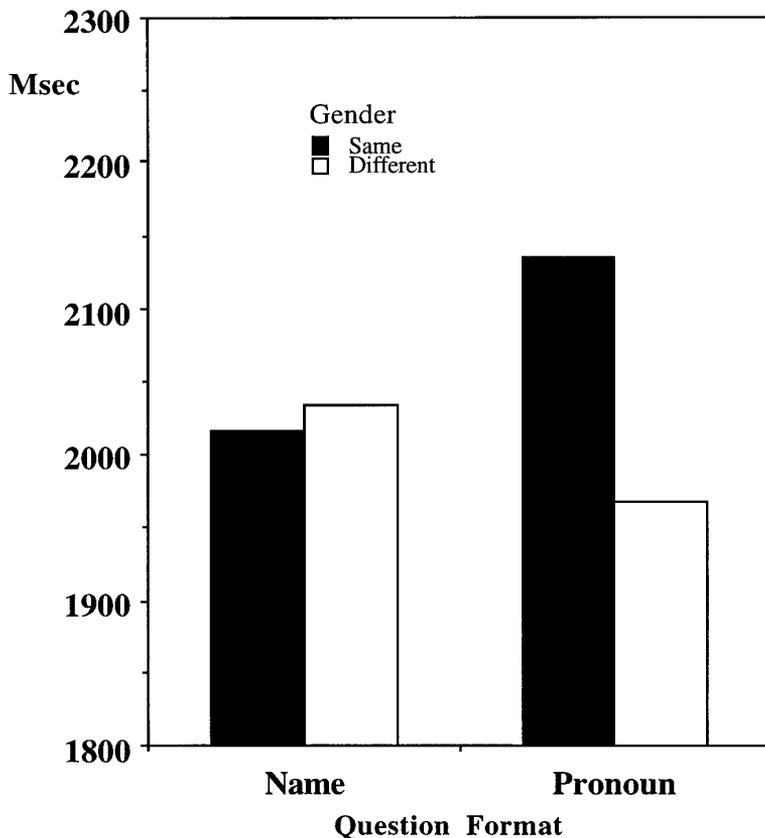
FIG. 1. Experiment 1: Verification latencies of correct responses, as a function of the format of the question and the match between the gender of the actors in the privileged sentence and the question.

a significant effect, $F1(1,47) = 4.37, p < .05$, $MSE = .0298, F2(1,15) = 6.15, p < .03, MSE = .0070$. Reflecting the fact that the Gender effect occurred only for pronouns, the interaction was significant, $F1(1,47) = 4.72, p < .04$, $MSE = .0223, F2(1,15) = 6.64, p < .03, MSE = .0053$.

Both the reaction time and error-rate data yielded the pattern predicted by the Unrestricted Search hypothesis. Addressees made more errors and took longer to respond to the question ''Did he break a leg?'' when their privileged knowledge provided them with a potential referent for the pronoun, i.e., when the sentence that was inaccessible to the speaker was about a male character rather than a female character. This strongly supports the assumptions of the Unrestricted Search hy-

pothesis: It suggests a search for referents that is not restricted to entities that are accessible to the *speaker*. This unrestricted search led to more errors and delayed correct responses when entities that were not accessible to the speaker were potential referents of the pronoun.

Experiment 1 demonstrated the interference predicted by the Unrestricted Search hypothesis during the resolution of pronouns—a form of definite reference. It also showed that the predicted effect occurs when the procedure guarantees that participants have ample opportunity to memorize the accessibility status of information. This demonstration should rule out an alternative explanation of the results which assumes that participants might have simply been confused. In addition, partici-
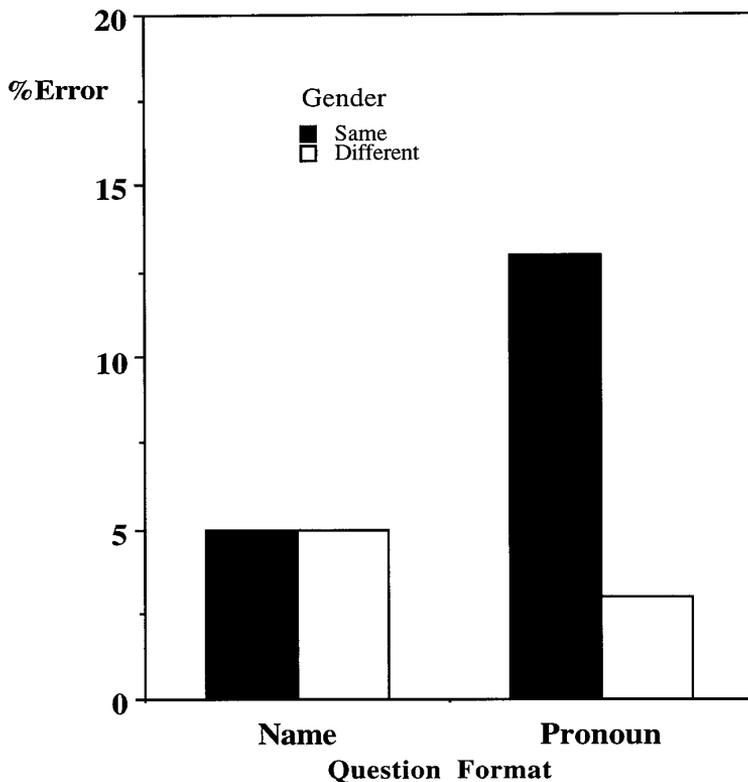
FIG. 2. Experiment 1: Error rates as a function of the format of the question and the match between the gender of the actors in the privileged sentence and the question.

pants in this experiment believed that they were addressed by a real speaker whom they just met and was presumably talking to them via an intercom. This feature of the experiment contributes to its generalizability.

Recall the real-life analog of our lab setup. The reason the Unrestricted Search hypothesis predicts that Boris would experience interference when his wife asked, ''Is she asleep?'' is that he was thinking about his secret lover right before she asked the question. His lover was therefore prominent in his mind and was considered as a candidate for the pronoun ''she.'' Unlike Boris, who had no idea that his wife would burst in on his private thought, our participants had sufficient time to prepare for their interlocutor's question. For example, they could have foregrounded the relevant information and backgrounded the privileged information in preparation for the impending in-

teraction with their partner. Apparently, they did not do that.

We can now describe what might have been going through Boris' mind when his wife asked him, ''Is she asleep?'' Most likely, Boris experienced interference in resolving the pronoun ''she.'' Given that Boris' lover was accessible to him, and that she is a potential referent of the pronoun, our data suggest that Boris would entertain an interpretation which assigns his lover to the pronoun. This should occur even though his wife could not have been referring to his lover. If indeed Boris creates such an interpretation, then he should behave just like our addressees. He might revise the interpretation to include only referents which are part of his common ground with his wife. In that case, he will assign their daughter to the pronoun ''she'' and answer, ''No, she is not asleep.'' The correct response will be

delayed because of the need for a revision (we will discuss the possible nature of such a revision process later in the paper). Alternatively, his interpretation might be compelling enough. In this case, he will do what our participants did when they made an error, and respond, ''Yes, she probably is.''

EXPERIMENT 2: DEMONSTRATIVE REFERENCE IN REAL CONVERSATION

The main goal of Experiment 2 is to generalize the findings of Experiment 1 to the context of an actual conversation. In addition, it is possible that addressees in Experiment 1 used the privileged information only because the first sentence was not established as the mutually known topic. In Experiment 2 the target referent is clearly established in advance as the common topic of the conversation. If referents which are not part of that common topic but are privileged to the addressee are still considered as intended referents under these conditions, it would be strong evidence against the Restricted Search hypothesis.

Experiment 2 investigates the comprehension of demonstrative reference, partly because the role of common ground in understanding this particular form of definite reference has been experimentally tested in the past (Clark, Schreuder, & Buttrick, 1993). Clark et al. consider the role of salience in the interpretation of demonstrative reference, and illustrate their theory with the following example. Suppose that Julia nods toward several men and says to Ken, ''That man is my neighbor.'' How would Ken interpret the demonstrative reference ''that man?'' As Clark et al. propose, each of these men can be salient on different grounds. One can be perceptually salient (e.g., very fat), another can be salient because he is leading the group and so on. Clark et al. propose that Ken would pick the man who is

> . . . most salient not on *general* grounds, but against their *particular* common ground. He was to select the bald man even if the tallest man, the

midget, or the winner would be most salient on general grounds. As the principle of optimal design dictates, the only information he should consult is their common ground.'' (p. 247)

In other words, according to the Restricted Search hypothesis, salience plays a role only among common ground entities.

In contrast, the Unrestricted Search hypothesis assumes that salient referents will be considered regardless of whether or not they are in common ground. This is because the hypothesis assumes a search for referents which is not restricted to common ground. Such a search should immediately consider salient referents. Clark, Schreuder, and Buttrick's (1983) experiments cannot distinguish between the two hypotheses because their experiments confound salience with common ground (see Keysar, 1997, for detailed discussion). The critical test is not whether salient referents in common ground would be selected, but whether salient referents which are *not* in common ground would be considered. The Restricted Search hypothesis assumes that the answer is no, while the Unrestricted Search hypothesis assumes the answer is yes. Experiment 2 tested these assumptions.

The rationale of the second experiment is analogous to the following situation. Suppose you are attending a gallery tour of a Matisse exhibit in an art museum. The group is standing around the painting ''Portrait de Marguerite endormie,'' and the guide is talking about it. Your eyes wander about and you find yourself looking at the picture behind the guide, ''Lorette a la veste rouge.'' At that moment, you hear the guide say ''This woman is Matisse's daughter.'' Who would you take to be the referent of ''this woman,'' Marguerite or Lorette? You know that the guide is talking about Marguerite and that his back is turned to the painting of Lorette. So, obviously you would eventually pick Marguerite as the referent of ''this woman.'' The Restricted Search hypothesis predicts that you will search for referents only among entities which are common; therefore, you will not consider Lorette who is clearly not part of the common ground with the guide. Yet, the Unrestricted Search

hypothesis predicts that because Lorette is perceptually available to you and is a potential referent, you will consider her as the referent. Had we been able to measure your comprehension processes in the example above, then we should find temporary interference when you happen to consider Lorette's picture compared to one that does not have a potential referent, as in ''Paysage d'oliveiers,'' a picture of olive trees.

It is difficult to test for this type of interference with an on-line measure without disrupting the flow of conversation. A relatively unobtrusive methodology is described by Tanenhaus and his colleagues (Eberhard, Spivey - Knowlton, Sedivy, & Tanenhaus, 1995; Tanenhaus, Spivey-Knowlton, Eberhard, & Sedivy, 1995), who tracked addressees' eye movements while they followed instructions. Tanenhaus and his colleagues established that when addressees interpret a referring expression, their eye gaze will fixate immediately on a potential referent object. Therefore, one can use eye fixations as a sensitive measure of on-line comprehension. In the context of our Matisse example, as soon as you hear the guide say, ''This woman is Matisse's daughter,'' you might move your gaze from Lorette's picture back to Marguerite's picture which is the known topic of discussion. Yet, if the Unrestricted Search hypothesis is correct, then the launch of that saccade should be delayed because you are actually considering Lorette as the referent. In Experiment 2, we created a situation which is analogous to this example in order to test for such a delay in the context of an actual conversation.

Method

Participants. Thirty native English speakers contributed data to this experiment. All were University of Chicago college students. None had participated in a similar experiment before; they were paid for their time. We replaced four participants due to missing data (i.e., inadequate calibration or experimenter errors) and two participants because they guessed that the ''other'' participant was a confederate.

Equipment. In order to track participants' eye movements, we used an Applied Science Laboratories head-mounted video camera connected to an eye-tracking unit. The camera lens was mounted on a headband that fit around the participant's head. The head-mounted equipment was lightweight (about 9.5 oz) and did not restrict his or her ability to move naturally. It provided information about eye movements with respect to the head, and a magnetic head tracker provided information about movements of the head. The combination of these pieces of information determine the location of the eye fixation. An additional camera was mounted on a post and recorded the pictures the participant was observing. Our data were a video image of the pictures with a superimposed cross hair representing the location of the eye fixation. This allowed us to determine where the participant's eye was fixating in any video frame, at a rate of 30 frames per second.

Experimental setup and procedure. The experiment was a variation of the ''referential communication task'' (e.g., Glucksberg, Krauss, & Higgins, 1975; Krauss & Glucksberg, 1977; Krauss & Weinheimer, 1964), where two participants converse about a set of objects. In each trial, one participant saw a set of four pictures mounted on a 20 by 30 inch cardboard sheet. The labeled pictures were situated at each corner of the board, 6 inches from the top or the bottom and 9 inches from the side. During the trial, the cardboard was placed vertically between the two participants to prevent them from seeing each other. The other participant, who was actually a confederate, received an outline of *one* of those four pictures, several color markers, and a crayon set. Their goal was to communicate so that the confederate would be able to add to the outline some missing details (e.g., color) and make it look more like the complete picture. We will refer to the confederate as ''the artist'' and to the real participant as ''the helper.'' Each trial was limited to one minute, and we asked the participants to do their best within that time frame.

The experimenter assigned roles to the participants in a seemingly random manner, mak-

ing sure the participant was always the helper, and then explained the task. The pair first practiced with one picture set and then reversed roles to ensure that the helper understood the artist's task. Before the role reversal, the experimenter introduced the eye-tracking equipment and explained our interest in measuring pupil diameter during conversation. No mention of eye movements was made so that participants did not become overly conscious of the way their eyes move.

After calibrating the equipment, the experimenter provided a cover story for the main manipulation of the experiment. As in the Matisse example, in which the picture of Marguerite is the topic of conversation, the target picture in the experiment was mutually established as the topic. In order to create the analogous case to the picture of Lorette (i.e., the picture not in common ground), we needed to create a situation where the helper would be looking away from the target picture and at another picture. To motivate this, the experimenter explained that ''the eye camera is very sensitive and can move out of alignment very easily,'' and that therefore he needed to check calibration throughout the experiment. To do this without interrupting the main task, he explained, he would find natural pauses, when the artist is busy filling in the information and play a sound file that would give calibration directions to the helper. The experimenter explained that to avoid interrupting the artist, the directions would be played through an earphone which was placed on one of the helper's ears. The calibration instructions included (1) a directive to look at one of the other three pictures and (2) a question about that picture. The goal of the question was to make sure that the helper considered the picture for a minimum amount of time before returning his or her gaze to the target picture. The helper was asked to answer the question in his or her mind and then to look back at the target picture. The experimenter emphasized that only the helper would be able to hear the recorded question.

Here is an example of one such trial. The helper had four pictures labeled ''plane,'' ''bowls,'' ''purse,'' and ''bird''; the artist received an outline of the target picture, the plane. Then the helper provided some information about the color of the cockpit and there was a pause while the artist was coloring in the details. At that point, the helper heard the calibration instructions ''Look at the bird. Is its beak long?'' and was supposed to mentally answer the question. Unbeknownst to the helper, the artist had a hidden earphone and was listening in on the calibration instructions. This allowed her to ask her next question immediately following those instructions. We refer to the artist's question that followed the calibration instructions as the critical question. In this example, she asked, ''Its wings, what color are they?'' This critical question is analogous to hearing the tour guide say, ''This woman is Matisse's daughter'' when you are looking at the picture behind him. So the question we asked was: What would the helper interpret as the referent of ''its wings?''

The Unrestricted Search hypothesis assumes that the search for referents is not restricted to common ground and would therefore take ''its wings'' to refer to the bird's wings. This hypothesis predicts interference in arriving at the correct interpretation of the critical question, which should appear as a delay to launch a saccade from the bird to the airplane. To test this prediction, we included a baseline condition, consisting of a calibration picture that had no potential referent for the artist's question. In such a baseline condition, then, we expect no interference. In this example, we compared the launch time from the word ''wings'' when the calibration picture was the bird to when it was a picture of a woman. In contrast to the Unrestricted Search hypothesis, the Restricted Search hypothesis assumes that only referents in common ground would be considered. Therefore, it does not predict a difference between the time to launch the saccade when the participant is calibrating to a picture of a bird compared to that of a woman.

Throughout the experimental session, we went to great lengths to ensure that the helper did not suspect the confederate because we

thought it important that the helper participate in the conversation as naturally as possible. During debriefing, we verified that we were successful by gradually probing the participants about their ''team work,'' and finally we explained that sometimes we must include confederates in our experiments to control for certain factors. At that point, we offered participants bonus pay if they correctly determined whether or not the artist was a confederate. All the participants guessed that the artist was not a confederate, except for two, who were replaced. The experiment lasted less than an hour.

Materials. The experiment had 12 items, each displayed on a different board. The items were modeled after the airplane/bird example above. Each item had a target picture about which the participants conversed and another picture which was ostensibly used for calibration; we will call this picture the ''calibration picture.'' We recorded calibration instructions to digital sound files. For each calibration picture, we recorded instructions consisting of two utterances, starting with ''look at the [picture label]'' which directed the helper to the picture, followed by a question about the calibration picture. In order to allow the helper to easily distinguish between the recording and the confederate's instructions, we recorded a female research assistant whose voice was clearly different from the voice of the confederate and the experimenter. To make sure the helpers considered the calibration questions, the experimenter quizzed them informally three times during the experiment. In the experimental condition, the calibration picture included a potential referent for the critical question (e.g., a picture of a bird). In the baseline condition, the calibration picture did not have a potential referent (e.g., a picture of a woman). Of course, the target picture always had a referent for the critical question.

The materials were divided into two versions, with each version including all 12 items, half in the experimental condition and half in the baseline condition, with the two versions using the mirror-image distribution of conditions over items. Each participant received only one of the versions. Late in the experiment, we realized that one item was constructed incorrectly (i.e., the calibration and target pictures were switched). We therefore do not include the data from this item in the analysis. We added three fillers by having the participants converse over a second picture on three of the experimental boards, after they had finished with the experimental item. Each participant had one practice trial, one role reversal, two warm-up trials followed by 12 items, and three fillers which were presented in random order. We used four different random orders counterbalanced with the experimental version and the helper's gender.

Design and predictions. The experiment had one within-subject independent variable; the calibration picture either had a potential referent or not. The Unrestricted Search hypothesis predicts that when the calibration picture has a potential referent for the critical question, helpers would take longer to launch their eyes back to the target picture than when the calibration picture does not have a potential referent (i.e., the baseline condition). The Restricted Search hypothesis assumes that only referents in common ground would be considered. It assumes that participants would look for referents only in the mutually established target picture.

Results and Discussion

A research assistant who was blind to the hypothesis coded the video tapes after an initial reliability test revealed close to perfect agreement with another coder. For each item, the coder determined the number of video frames that elapsed from the beginning of the referring expression in the artist's critical question (e.g., ''wings'' in ''Its wings, what color are they?'') to the launching of the eye toward the target picture. The video frames were converted to latency in ms. Two outlier latency points were not included in the analysis, neither were trials when the helper was not looking at the calibration picture when the critical question started, because we could not test the hypothesis in these cases.

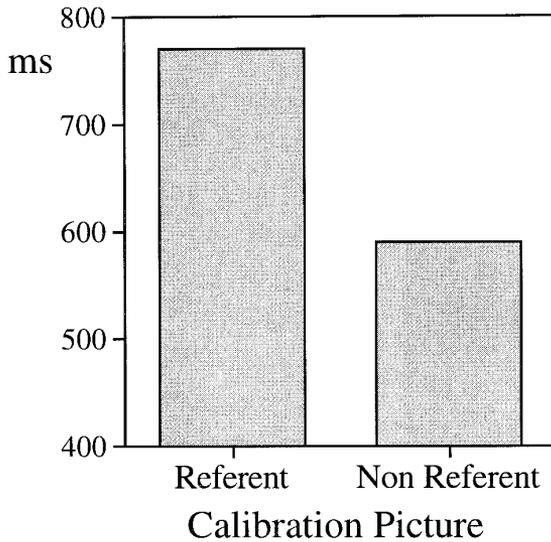As illustrated in Fig. 3, the pattern of sac-

FIG. 3. Experiment 2: Latency to launch a saccade back toward the target picture, from the beginning of the identifying noun phrase in the critical question, by whether or not the calibration picture had a potential referent.

cade launch latencies supported the Unrestricted Search hypothesis. When the calibration picture included a potential referent, the saccade launch was delayed an average of 180 ms. In general, participants differed vastly in initiating a launch. To eliminate scaling effects that might result from such differences between participants, we converted the latencies for each participant into z scores. For each participant, we averaged the scores separately for items which appeared in the experimental and baseline conditions and submitted these means to a paired t test. The analysis was significant, $t(29) = 3.44$, $p < .01$. Similarly, we averaged the scores for each item separately for participants who received the item in the experimental and baseline conditions, paired t test marginally significant, $t(10) = 1.97$, $p = .076$.

Given that the confederate artist was aware of the condition in each trial because she listened in on the calibration instructions, we were concerned that she might have contributed to the effect in some way. The most plausible way the artist could have influenced the launch of the helper's eye might be by unconsciously varying the synchronization between her question and the calibration instructions. Suppose that the artist systematically delayed producing her critical question in the baseline condition compared to the experimental condition. The longer the delay, the faster the helper might be to launch the eye back to the target picture. For example, it might be that a delayed critical question gives the helper more time to prepare to return to the target picture, and that in contrast, when a question comes immediately on the heels of the calibration instructions the helper is cognitively loaded, still trying to answer the recorded question in his or her mind, consequently taking longer to launch the saccade. We tested this possibility by measuring the latency from the end of the recorded instructions to the beginning of the artist's question. On average, the latencies for the experimental and baseline conditions were virtually identical (Means = 317 and 333 ms, respectively), $t < 1$. We can therefore rule out this possible artifactual explanation of our results.

GENERAL DISCUSSION

Clark and Carlson (1981, p. 328) foreshadowed the Unrestricted Search hypothesis

when they suggested that a comprehension process that is not restricted to common ground would ''systematically misinterpret'' definite reference. The data from Experiments 1 and 2 reveal just such systematic misinterpretations. This pattern of interference supports the Unrestricted Search hypothesis and rejects the Restricted Search hypothesis.

One might ask, how often do addressees experience such interference? According to the Unrestricted Search hypothesis, this should depend on the extent to which there is a discrepancy between the perspectives of the speaker and the addressee. It predicts interference only when perspectives diverge and when the addressee has privileged knowledge of potential referents, as in the Boris and Matisse examples. If perspectives rarely diverge, such interference should be uncommon.

What is the Role That Common Ground Plays?

Our experiments reject the accepted answer to this question and support the idea that the meta-knowledge that information is mutual does not restrict the comprehension of definite reference. This conclusion should constrain any process model of comprehension. However, the experiments do not tell us exactly what role common ground does play in comprehension. Process models could provide a variety of answers to this question, ranging from the assumption that mutual knowledge provides ''some'' constraint on the search to the assumption that mutual knowledge plays no role in the comprehension process. We propose a Perspective Adjustment model that assumes that common ground plays a corrective role in comprehension.

The Perspective Adjustment Model

Our model assumes the operation of two processes during comprehension: A fast, unrestricted search that interprets the definite reference by assigning a referent with no regard to mutual knowledge. This process is coupled with a monitoring and adjustment process that is sensitive to considerations of common ground. It uses the meta-knowledge that an entity is mutually known and attempts to correct violations of common ground. In contrast to the unrestricted search, the adjustment process is relatively slow—mainly because it activates higher level, meta-knowledge memory structures. The model assumes that the two processes proceed not in a strict serial fashion but instead in cascades (McClelland, 1979). The two processes operate continuously and the search process need not be fully completed before information is monitored for violation of common ground. Yet because the search is much faster than the monitoring process, the monitoring process is rarely able to preempt or constrain it.

It is interesting to note that the average interference effect was very similar in the two experiments. In Experiment 1 the interference delayed correct responses by a mean of 170 ms and in Experiment 2 it delayed launching a saccade by 180 ms. This relatively small effect suggests that the search process might not have been fully completed before partial products of the search were monitored. This is consistent with the possibility that the two processes of the Perspective Adjustment model operate in cascade.

According to the Perspective Adjustment model, then, the results of our experiments reflect the participants' need to correct a mistaken interpretation. When the perspectives of the two interlocutors diverged, the addressee's quick unrestricted search picked the wrong referent, which in turn required a correction. The slow correction is reflected in the interference pattern in both experiments. So, the Perspective Adjustment model provides a new answer to our question: The role of mutual knowledge is only to correct interpretation errors.

Next, we will consider alternative models that could answer the question differently. The original theories of mutual knowledge have been about its role during conversation, and our Perspective Adjustment model follows this tradition. In contrast, most of the research in psycholinguistics concerns memory processes during reading. The study of reading comprehension could explore the role of com-

mon ground, but only indirectly because readers are typically ''overhearers'' or ''side participants'' and not addressees. Consequently, experiments on readers typically investigate how the participants perceive the common ground among story protagonists; this is different from our experiments, which directly evaluate the role that common ground plays for addressees. It is therefore not obvious that one can readily generalize from research on side participants to actual addressees. Moreover, given that most theories of comprehension do not attempt to spell out a role for mutual knowledge, we will only conjecture about possible extensions to the study of conversation.

Extrapolation from Memory-Based Models of Reading Comprehension

Some memory-based models can explain our results. Greene, Gerrig, McKoon, and Ratcliff (1994) argue that characters in text serve as memory cues to information associated with them. The associated information is facilitated and becomes readily available for use. They report experiments that show that once protagonists in a story are reunited, information associated with them becomes available. This information can then be used to quickly understand a conversation between the characters who allude to that information. For example, in one story Jane tells Gloria she is going to dinner with her cousin. When she reunites with Gloria, the concept ''cousin'' becomes readily available and can facilitate the comprehension of the pronoun ''she'' as referring to the cousin if Jane says, ''She was very boring.''

Greene et al. interpret this ''reunion effect'' with respect to common ground. They argue that when characters in a story unite, information which is in their common ground is facilitated. Not just any information which is associated with the characters is cued by the reunion, but specifically mutually known information: ''. . . to the extent that Gloria and Jane's earlier conversation supports the cousin as mutually known, the cousin will be restored to the reader's focus of attention . . .'' (p. 514). The answer that this model

provides for our question about the role of common ground is: Knowing that information is in common ground selectively affects facilitation of information from long-term memory.

Greene et al.'s experiments do not actually warrant such conclusion about the role of common ground because they confound common ground information with information which is associated with the characters (for a detailed discussion, see Keysar, 1997). We do not know from these experiments whether the reunion facilitated the concept ''cousin'' because it was part of common ground, or because it was merely associated with the characters. Lea, Mason, Albrecht, Birch, and Myers (in press) demonstrated that the latter is correct—the reunion simply facilitates information which is associated with characters. They showed that ''cousin'' is facilitated regardless of whether it is part of common ground or just associated with Jane. These experiments show that the common ground does not play a role in the cued activation of entities as suggested by Greene et al.

Indeed, McKoon, Gerrig, and Greene (1996) report very similar reunion effects and do not explain them in terms of common ground. Instead, the model in McKoon et al. explains the reunion effect as a result of facilitation of associated information in general. The model as it was described in the McKoon et al. paper does not answer our question regarding the role of common ground in comprehension because it does not spell out a role for mutual knowledge in the comprehension process. Future extensions might include a role for the meta-knowledge that information is mutual.

Another possible memory-based model that could be extended to account for our results comes from the work of Gernsbacher and her colleagues, who demonstrate that relevant information is enhanced during reading while irrelevant information is actively suppressed (e.g., Gernsbacher, 1989, 1990; Gernsbacher & Faust, 1991). Perhaps by suppressing activated information which is not mutually known, addressees are able to arrive at the intended referent. A model of this type would

answer the question in focus as follows: The role of common ground is to suppress information which is not common.

Processing and Optimality

Clark and Carlson (1981) proposed that "the comprehension process must keep track of common ground, and its performance will be optimal if it limits its access to that common ground. Whether its design is actually optimal in this respect is a question that can only be answered empirically." (p. 328). Clark, Schreuder, and Buttrick (1983) argued that the empirical answer is positive, that the comprehension system is designed to search for referents only among entities in common ground. We challenge this conclusion and propose that the design of the system is different. It is not designed to search among entities in common ground, but instead to use common ground only to correct errors.

We also challenge Clark and Carlson's assumption about optimality. It is not obvious that a comprehension system which limits its access to common ground would perform optimally. If optimal operation means error-free, then the Perspective Adjustment model is not optimal, but if one takes into account the "cost" of restricting the search to common ground information, the Perspective Adjustment model might very well be optimal. Using higher level common-ground information to restrict the search for referents might tax the cognitive resources of the system and might not be worth the advantage of avoiding the occasional error. It might turn out, then, that in terms of cost and benefit the Perspective Adjustment model is optimal. It might also turn out that the comprehension system adapts to changing circumstances so that it restricts access to common ground when it is cost effective but it does not restrict access when the cost outweighs the benefit. But these questions can only be answered in the future.

APPENDIX: IMPORTANT EXCERPTS FROM THE INSTRUCTIONS TO PARTICIPANTS IN EXPERIMENT 1.

1. The written instructions explained the role of the second sentence in the following way:

"One of our goals is to see how people keep two different pieces of information in mind during a conversation. To that end, we will give you an additional sentence which you will have to keep in mind while you answer your partner's question."

2. This is the section from the instructions which stressed the fact that the second sentence is inaccessible to the speaker and is irrelevant to the scenario. The bold type face appeared in the original instructions.

"In each trial, the **first** sentence we give you will always be what you need to know to answer your partner's question. It is important to note that the second sentence is **not** to be used to answer your partner. It is irrelevant to your partner because he or she knows nothing about the person mentioned in that second sentence. For example, your partner knows nothing about Marla; he or she received a story about Rachel and Joe. Therefore, the sentence about Marla is irrelevant to the answer and cannot help you give the correct answer to your partner. Therefore, you should not use it when answering your partner's question. However, it is crucial that you read and remember it, because you might eventually have to recall it."

REFERENCES

Caramazza, A., Grober, E., Garvey, C., & Yates, J. (1977). Comprehension of anaphoric pronouns. *Journal of Verbal Learning and Verbal Behavior,* **16,** 601–610.

Chafe, W. L. (1976). Givenness, contrastiveness, definiteness, subjects, topics, and point of view. In C. N. Li (Ed.), *Subject and topic.* New York: Academic Press.

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

Clark, H. H., & Carlson, T. B. (1981). Context for comprehension. In J. Long & A. Baddeley (Eds.), *Attention and performance IX.* Hillsdale, NJ: Erlbaum.

Clark, H. H., & Carlson, T. B. (1982). Speech acts and hearers beliefs. In N. V. Smith (Ed.), *Mutual knowledge.* London: Academic Press.

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

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., & Schaefer, E. F. (1989). Contributing to discourse. *Cognitive Science,* **13,** 259–294.

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., McKoon, G., & Ratcliff, R. (1983). The activation of antecedent information during the processing of anaphoric reference in reading. *Journal of Verbal Learning and Verbal Behavior,* **22,** 121–132.

Donnellan, K. (1966). Reference and definite descriptions. *Philosophical Review,* **75,** 281–304.

Eberhard, K. M., Spivey-Knowlton, M. J., Sedivy, J. C., & Tanenhaus, M. K. (1995). Eye movements as a window into real time spoken language comprehension in natural contexts. *Journal of Psycholinguistics Research,* **24,** 121–135.

Ehrlich, K. (1980). Comprehension of pronouns. *Quarterly Journal of Experimental Psychology,* **32,** 247–255.

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,** 203–219.

Gernsbacher, M. A. (1989). Mechanisms that improve referential access. *Cognition,* **32,** 99–156.

Gernsbacher, M. A. (1990). *Language comprehension as structure building.* Hillsdale, NJ: Erlbaum.

Gernsbacher, M. A., & Faust, M. E. (1991). The mechanism of suppression: A component of general comprehension skill. *Journal of Experimental Psychology: Learning, Memory, and Cognition,* **17,** 245–262.

Gerrig, R. J., & Littman, M. L. (1990). Disambiguation by community membership. *Memory and Cognition,* **18,** 331–338.

Gibbs, R. W., Mueller, R. A. G., & Cox, R. W. (1988). Common ground in asking and understanding questions. *Language and Speech,* **31,** 321–335.

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

Glucksberg, S., Kreuz, R. J., & Rho, S. H. (1986). Context can constrain lexical access: Implications for models of language comprehension. *Journal of Experimental Psychology: Learning, Memory, and Cognition,* **12,** 323–335.

Greene, S. B., Gerrig, R. J., McKoon, G., & Ratcliff, R. (1994). Unheralded pronouns and management by common ground. *Journal of Memory and Language,* **33,** 511–526.

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

Hogaboam, T. W., & Perfetti, C. A. (1975). Lexical ambiguity and sentence comprehension. *Journal of Verbal Learning and Verbal Behavior,* **14,** 265–274.

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

Johnson-Laird, P. N. (1982). Mutual ignorance: Comments on Clark and Carlson's paper. In N. V. Smith (Ed.), *Mutual knowledge.* London: Academic Press.

Johnson-Laird, P. N., & Garnham, A. (1980). Descriptions and discourse models. *Linguistics and Philosophy,* **3,** 371–393.

Keysar, B. (1997). Unconfounding common ground. *Discourse Processes,* **24,** 253–270.

Keysar, B. (in press). Language users as problem solvers: Just what ambiguity problem do they solve? In S. R. Fussell & R. J. Kreuz (Eds.), *Social and cognitive psychological approaches to interpersonal communication.* Hillsdale, NJ: Erlbaum.

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

Keysar, B., Barr, D. J., & Horton, W. S. (in press). The egocentric basis of language use: Insights from a processing approach. *Current Directions in Psychological Sciences.*

Kintsch, W., & Mross, E. F. (1985). Context effects in word identification. *Journal of Memory and Language,* **24,** 336–349.

Krauss, R. M., & Fussell, S. R. (1996). Social psychological models of interpersonal communication. In E. T. Higgins & A. Kruglanski (Eds.), *Social psychology: Handbook of basic principles.* New York: Guilford Press.

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.

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

Lea, R. B., Mason, R. A., Albrecht, J. E., Birch, S. L., & Myers, J. L. (in press). Who knows what about whom: What role does common ground play in accessing distant information? *Journal of Memory and Language.*

Malt, B. C. (1985). The role of discourse structure in understanding anaphora. *Journal of Memory and Language,* **24,** 271–289.

McClelland, J. L. (1979). On the time relations of mental processes: An examination of systems of processes in cascade. *Psychological Review,* **86,** 287–330.

McKoon, G., Gerrig, R. J., & Greene, S. B. (1996). Pronoun resolution without pronouns: Some conse-

quences of memory-based text processing. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 22,* 919–932.

McKoon, G., & Ratcliff, R. (1992). Inference during reading. *Psychological Review, 99,* 440–466.

McKoon, G., Ward, G., & Ratcliff, R. (1993). Morphosyntactic and pragmatic factors affecting the accessibility of discourse entities. *Journal of Memory and Language, 32,* 56–75.

Mueller-Lust, R. A., & Gibbs, R. W. (1991). Inferring the interpretation of attributive and referential definite descriptions. *Discourse Processes, 14,* 107–131.

Murphy, G. L. (1985). Process of understanding anaphora. *Journal of Memory and Language, 24,* 290–303.

Onifer, W., & Swinney, D. A. (1981). Accessing lexical ambiguities during sentence comprehension: Effects of frequency of meaning and contextual bias. *Memory and Cognition, 9,* 225–236.

Ortony, A., & Anderson, R. C. (1977). Definite descriptions and semantic memory. *Cognitive Science, 1,* 74–83.

Rinck, M., & Bower, G. (1995). Anaphora resolution and the focus of attention in situation models. *Journal of Memory and Language, 34,* 110–131.

Sadock, J. M. (1978). On testing for conversational implicature. In P. Cole (Ed.), *Syntax and semantics 9: Pragmatics.* New York: Academic Press.

Simpson, G. B. (1981). Meaning dominance and semantic context in the processing of lexical ambiguity. *Journal of Verbal Learning and Verbal Behavior, 20,* 120–136.

Sperber, D. (1982). Comments on Clark and Carlson's paper. In N. Smith (Ed.), *Mutual knowledge* (pp. 46–51). London: Academic Press.

Sperber, D., & Wilson, D. (1982). Mutual knowledge and relevance in theories of comprehension. In N. Smith (Ed.), *Mutual knowledge.* London: Academic Press.

Sperber, D., & Wilson, D. (1986). *Relevance: Communication and cognition.* Cambridge, MA: Harvard University Press.

Sperber, D., & Wilson, D. (1987). Precis of Relevance: Communication and cognition. *Behavioral and Brain Sciences, 10,* 697–754.

Swinney, D. A. (1979). Lexical access during sentence comprehension: (Re)consideration of context effects. *Journal of Verbal Learning and Verbal Behavior, 18,* 645–659.

Tabossi, P., Colombo, L., & Job, R. (1987). Accessing lexical ambiguity: Effects of context and dominance. *Psychological Research, 49,* 161–167.

Tanenhaus, M. K., & Lucas, M. M. (1987). Context effects in lexical processing. *Cognition, 25,* 213–234.

Tanenhaus, M. K., Spivey-Knowlton, M. J., Eberhard, K., & Sedivy, J. C. (1995). Integration of visual and linguistic information in spoken language comprehension. *Science, 268,* 1632–1634.

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