Unconfounding Common Ground

BOAZ KEYSAR
The University of Chicago

I propose that research on the role of common ground in language use is often insensitive to considerations of parsimony. By ignoring parsimony, our experiments might fall short of addressing common ground and instead the resulting data support a simpler theory which is subsumed within the theory of common ground. In order to address this inherent conceptual confound, this paper proposes the subsuming theory criterion for experiments on common ground and mutual knowledge. In practical terms, to demonstrate an effect of common ground the design must keep common information constant and only vary whether or not it is common. The paper demonstrates how consistent use of this necessary criterion will allow us to make stronger claims about when common ground plays a role and when it does not. To illustrate how experiments can be designed to satisfy the criterion, the paper evaluates some earlier studies which do and don't satisfy it. In general, the study of language use could benefit from tightening the rules of evidence, allowing us to draw stronger conclusions about central concepts such as common ground.

The relationship between data and theory in science is fairly complex. Rarely do experiments provide data which uniquely support a particular theory, but instead data are typically interpretable in more than one way. In order to improve our ability to interpret data and select among potential explanatory theories, we rely on accepted "rules of evidence." One of the most central rules of evidence is the principle of parsimony. If an experiment provides evidence which could support two alternative theories, and one theory is clearly more parsimonious than the other, then in general one should consider the experiment as supporting the more parsimonious theory. In other words, to the extent that experiments affect our beliefs, the experiment *should not* affect the strength of our belief in the less parsimonious theory. While parsimony is taken for granted in the general practice of

Direct all correspondence to: Boaz Keysar, Department of Psychology, The University of Chicago, 5848 South University Avenue, Chicago, IL 60637 <boaz@ccp.uchicago.edu>.

Discourse Processes, 24, 253–270 (1997)
ISSN 0163-853X

science, the experimental study of language use tends to be less sensitive to considerations of parsimony. In particular, my argument is that the experimental study of the theory of common ground does not consistently apply the principle of parsimony. I will suggest that a more consistent use of this principle and general rules of evidence would yield a more direct link between our experiments and our understanding of the role that common ground plays in communication.

PARSIMONY AND PERFECTLY SUBSUMED THEORIES

In order to apply the principle of parsimony one needs a way to decide which theory is simpler. One possibility is to consider the theory which postulates fewer assumptions as more parsimonious. This might not be feasible when the theories and their assumptions are not comparable. For such occasions there might be an inherent difficulty in identifying the "basic" set of assumptions a theory makes, in deciding how to "parse" the assumptions and demonstrating that there are no "hidden" assumptions. This difficulty often renders the principle of parsimony practically inapplicable (Goodman, 1979). There are cases, however, when the principle of parsimony can be applied fairly easily. This is a description of what I would call "perfectly subsumed theories." A theory T which makes assumptions $A_1, A_2 \ldots A_n$, is perfectly subsumed within theory T' which makes assumptions $A_1, A_2 \ldots A_n, A_{n+1}$. The first theory is unambiguously more parsimonious than the second.

When an experiment provides data which can be explained by either the subsuming theory or the perfectly subsumed one, we should prefer the simpler theory, the one which is perfectly subsumed. Let me use a simple-minded example as a *rhetorical device* which will allow me to argue by analogy when I discuss theories of language use later in the paper. Consider the following situation. The Saint of Publications appears in a dream to Joseph, who is a struggling assistant professor. She introduces herself and informs him that whenever she blesses a good paper (and only a good paper) upon request, the paper is accepted for publication. Joseph takes this as a hypothesis to be tested, and in order to provide evidence for the theory, he writes a good paper, requests her blessings, and sends it off to a journal. Lo and behold, his paper is accepted for publication. He replicates this success many times and from then on becomes a follower of the Saint. Joseph's reasoning does not hold precisely because he violated the principle of parsimony. He interpreted the "data" as supporting a subsuming theory over a perfectly subsumed alternative theory. Consider the assumptions Joseph's theory will make in order to explain the findings (i.e., to explain the accepted publication):

Theory 1: The Subsuming Theory
(1) There is a paper P.
(2) The paper P is good.
(3) The Saint blessed P.
(4) Consequently, P will be accepted for publication.

This kind of "theory" subsumes a simpler "theory" which makes fewer assumptions:

Theory 2: The Perfectly Subsumed Theory
(1) There is a paper P.
(2) The paper P is good.
(3) Consequently, P will be accepted for publication.

Let us call the first the "blessed, good paper" theory and the second "the good paper" theory. It is obvious that the data (i.e., the fact that blessed good papers are accepted) do not distinguish the saint-driven theory from the simpler, good paper theory that it subsumes. The papers might have been accepted because they were good *and* blessed, or they might have been accepted simply because they were good: The data do not tell us which. Given that the good paper theory is simpler by definition, one should prefer it over the blessed, good paper theory for reasons of parsimony, at least until further evidence is provided.

It might be easy to see that the saint-driven theory is not supported by the data, especially if one does not believe in the saint's power over journal editors. In contrast, if one intuitively believes in the saint-driven theory to begin with, the insufficiency of the evidence may be more difficult to ascertain. Therefore, when one is dealing with a saint-type theory which is intuitively compelling, one could benefit from explicit criteria for evaluating the force of evidence. This might help us determine when the evidence only supports the subsumed theory and not the compelling yet subsuming theory. In this study I focus on one theory of language use which is highly compelling: the theory that common ground plays an important role in language use.

COMMON GROUND AND PERFECTLY SUBSUMED THEORIES

The Concept of Common Ground as Central in Theories of Language Use

Theories that assume that common ground is central to language use dominate the field of language use. To understand the ubiquity of this assumption consider the comprehension of something as prevalent as demonstrative reference. The

other day I asked a florist "How much is *that* flower?" while gesturing at a slew of different flowers. The use of "that flower" did not pick out a unique referent, but she immediately quoted the price for a stunning Bird of Paradise. This flower stood out in the otherwise mundane array of options. Current theories of language use postulate that the florist used common ground to interpret my utterance. The flower was mutually salient to us and therefore she assumed that I must have been referring to it. The concept of common ground was initially controversial (Sperber & Wilson, 1982; Johnson-Laird, 1982) but has become a central concepts in current psychological theories of language use. While this concept has been defined quite precisely (Clark & Marshall, 1981; Clark, Schreuder, & Buttrick, 1983; Gibbs, 1987), and a strong theoretical case has been made for the centrality of common ground in language use, I will argue that research in the field has not used a consistent standard of evidence to evaluate experimental data for common ground. In particular, the principle of parsimony has not been consistently applied. The goal of this study is to propose a clear criterion for experiments that purport to demonstrate the use of common ground.

What is Common Ground?

Common ground has been conceptualized as a type of shared information (Clark, 1992, p. 3; Schiffer, 1972; Stalnaker, 1978). But information that two people share is not always part of their common ground. The proposition p is shared information for persons A and B if A knows that p and B knows that p, but it is part of their common ground only if they also mutually know that p. I might know that Susan reviewed my paper, Susan obviously knows that she reviewed my paper, but this fact is part of our common ground only if she knows that I know, and I know that she knows, and so on. As Clark (1992) stated, "The common ground between Ann and Bob [...] is the sum of their mutual knowledge, mutual beliefs, and mutual suppositions" (p. 3).

Clark and Marshall (1981) describe three varieties of mutual knowledge: mutual knowledge as a function of community membership, physical co-presence and linguistic co-presence. If we both know that we are from Chicago then we can assume that we mutually know that the expression the "Kennedy Expressway" is a misnomer. By virtue of belonging to the community of Chicagoans, we mutually know that the Kennedy slogs along endlessly. In contrast to the community-membership basis for common ground which relies on long-term memory structures, physical and linguistic co-presence are a function of the immediate context. Clark and Marshall suggest that physical co-presence is "the strongest evidence of mutual knowledge that people are in general prepared to accept" (1981, p. 38). Clark (1996) replaces the notion of physical co-presence with the more psychologically grounded perceptual co-presence. If we are both sitting on that "express-

way," looking hopelessly at the cars around us, I will immediately understand your utterance "this is a parking lot." The traffic jam is part of our common ground because it is perceptually co-present. If I reply with "yes, but at least it is free parking" you will know what I am referring to by virtue of linguistic co-presence. Your preceding utterance, then, is our mutual linguistic context and serves as the basis for our common ground by virtue of linguistic co-presence. In general, once a certain piece of information is grounded by interlocutors (Clark & Brennan, 1991), they probably keep a record of it as part of a "model of the interlocutor": "The suggestion is that we carry around rather detailed models of people we know, especially of people we know well" (Clark & Marshall, p. 55). They argue that this model is employed by language users to infer the intention of speakers and to design utterances for their addressees.

It is reasonable to assume that people indeed have a model of the other, and that they update it precisely as Clark and Marshall (1981) describe. Indeed, Graesser, Bowers, Bayen, and Hu (in press) demonstrated that readers are quite good at keeping track of "who knows what," and Fussell and Krauss (1992) demonstrated that people accurately estimate what is known to members of different communities. But let us assume further that even when people do know who has access to what information, even when they have common ground information, they may or may not use it when they communicate. The question is, how does one demonstrate experimentally that common ground was actually *used* by an addressee to comprehend an utterance, or that it was consulted by a speaker to design an utterance? This paper is an attempt to establish minimally necessary conditions for such an empirical demonstration. Specifically, I will argue that a clear distinction between subsumed and subsuming theories will allow a more rigorous application of the principle of parsimony which in turn will improve the connection between our data and our theories.

When Can We Conclude that Common Ground Was Used?

Definition and Conditions. At first blush the answer seems trivial. If one wants to show that common ground was used by an interlocutor all one has to do is demonstrate that common ground existed and that the common information played a role in the production or comprehension of an utterance. Trivial as it seems, this assumption alone is insufficient because the use of the common information can be accounted for by a theory which is subsumed by the concept of common ground. To see this, consider the representation of mutual knowledge that Harman (1977) proposed, as phrased by Clark and Marshall (1981):

> A and B mutually know that p = def.
> (q) A and B know that p and that q.

This definition captures the notion of mutual knowledge, the essential element of common ground in the psychological literature. The definition of A and B mutually knowing p is that they each know it individually and that they both know that fact (q). So there are two elements to common ground—the piece of information (p) which is known to A and known to B, and the meta-knowledge which makes this information mutual (q).

Given this definition, we can outline what empirical research must demonstrate in order to infer the effect of common ground. If we would like to conclude that a listener used common ground in a particular case, the experiment must demonstrate that p is mutually known to the speaker and the listener and that the listener used p in order to interpret the speaker's utterance. But this would not be sufficient. The experiment must also demonstrate that the listener used p because it is part of the common ground with the speaker. So when I asked the florist "How much is that flower?" she probably used the salience of a particular flower (p) to interpret my utterance. The difficulty is to demonstrate that she identified the flower not only because it was salient to her but because she knew that it was mutually salient to us.

This analysis brings us to the core of the problem. In general, if A and B mutually know a certain piece of information (p), then by definition A knows that information. Now assume that A uses p to interpret B's utterance. She might use that information for either one of two reasons: because p is mutually known to her and to B, *or* simply because p is known or available to her. The flower example is a case of common ground by perceptual co-presence. Because the information is co-present it is by definition present for each individual. This suggests an inherent confound between the information being available to A and the meta-knowledge that it is mutually known. In this sense, a theory that invokes common ground is a special case of a theory which completely subsumes another: Information which is part of common ground is always information which is available to each individual alone. Therefore, a theory that argues that the information was used because it was common always subsumes a simpler theory that suggests that the information was used because it was available to the individual. The question is, how to distinguish between the theory which assumes a role for common ground and the perfectly-subsumed, more parsimonious theory which does not assume a role for common ground?

This conceptual confound poses the following problem for researchers who investigate the role that common ground plays in language use. If one would like to demonstrate that A used a certain piece of information because it was mutually known, one must also demonstrate that it was not used merely because it was available or accessible to A. In other words, one must show that information is used because it is co-present, not because it is simply present. Yet, given that it is

notoriously difficult to establish direct causality, a weaker criterion, the "subsuming theory" criterion, should be satisfied:

> *Show that information was used NOT because of factors which are perfectly subsumed by common ground. Or, show that subsumed theories are insufficient to explain the use of the information and that the subsuming theory is necessary.*

Suppose that an addressee, John, used *p* to interpret an utterance. The following two "theories" can account for the fact that John used *p*. One theory explains it by means of common ground:

Theory 1: The Subsuming Theory
(1) John believes *p*.
(2) The speaker believes *p*.
(3) John and the speaker mutually believe 1 and 2.

The simpler theory makes only assumptions 1 and 2, without assuming common ground:

Theory 2: The Perfectly Subsumed Theory
(1) John believes *p*.
(2) The speaker believes *p*.

In sum, if a listener used information which is common ground in order to interpret an utterance, it does not constitute evidence for the use of common ground. The reason is that it does not distinguish between a theory which assumes a role for common ground (the subsuming theory) and a theory which does not assume such a role (the perfectly-subsumed theory). Therefore, one must rule out the perfectly subsumed as a necessary condition for the data to be relevant for the theory of common ground.

APPLYING THE CRITERION

The subsuming theory criterion can be translated into practical terms to evaluate a particular research design. To show that the mutuality of the information, not the information itself, was necessary for comprehension, one would want to apply the subsuming theory criterion in the following way:

> *For an experiment to have potential to test the effect of common ground, keep the common information constant and manipulate its common ground status—i.e., manipulate whether or not it is common ground.*

By analogy to the saint example, in order to demonstrate the blessing effect predicted by the saint-driven theory one should keep the quality of the paper constant and manipulate the blessing, comparing a blessed good paper with a non blessed good paper. In the remainder of this section, I will demonstrate the use of the subsuming theory criterion by applying it to specific studies, starting with one detailed example and then discuss several others. I will first describe studies which do not satisfy the criterion and then some which do. Though the connection to common ground appears to be self-evident in some studies, if one applies the subsuming-theory criterion, one is forced to conclude that the data from these experiments are irrelevant to the issue of common ground. By applying this criterion, then, we should be able to distinguish between studies which provide data relevant to common ground and studies which do not.

Detailed Example: Do Readers use their Knowledge of Common Ground when They Resolve Unheralded Pronouns?

Consider a recently published paper which investigated the role of common ground during the comprehension of text (Greene, Gerrig, McKoon & Ratcliff, 1994). This particular paper is an excellent example because it presents a series of well-conducted experiments, it contributes an interesting set of findings, and the data seem to bear on the issue of common ground.

Design, Measures and Main Findings. The Greene et al.(1994) paper is concerned with the way readers use their knowledge of common ground between characters to manage the accessibility of information which in turn could be used to interpret an exchange between those characters. The basic idea is that when characters go in and out of the readers' focus their common ground follows suit and readers manage the accessibility of information by keeping track of common ground between characters which are in focus or present in the text.

Participants read scenarios which established certain concepts as part of the common ground between characters. Here is an example of such a scenario, where the concept cousin is established as part of the common ground between Jane and Gloria.

> Jane was dreading her dinner with her cousin, Marilyn. She complained loudly to her roommate Gloria. "Every time I go to dinner at my cousin's house I get sick." Gloria asked, "Why did you agree to go?" Jane said, "Because I'm too wimpy to say no." Jane went off to have dinner. [Concept-Present version begins.] When she arrived, Marilyn was just finishing the cooking. "You're in luck," she said, "we're having fried squid." Jane knew she was in for a wonderful evening. The two of them sat down to dinner. After dinner, they talked for a while and then Jane left. [Concept-Present version ends.] [Before Reunion test point.] Gloria was still up when Jane arrived home about midnight. [After Reunion test point.] Gloria asked

> Jane, "Did she poison you again?" [After Pronoun test point.] Jane chuckled and said, "We'll see in the morning" (p. 514).

In this "concept-present" version the reader "stays" with Jane when she goes to dinner and meets her cousin Marilyn, and therefore the common ground concept of cousin is present. In contrast, in the "concept-absent" version, the reader stays with Gloria when Jane leaves:

> Jane went off to have dinner. [Concept-Absent version begins.] Gloria decided to cook something nice for herself for dinner. "As long as I'm home alone," she thought, "I'll eat well." Gloria searched her refrigerator for ingredients. She found enough eggs to make a quiche. After dinner, she put the dishes in the dishwasher. [Concept-Absent version ends.] [Before Reunion test point.] Gloria was still up when Jane arrived home about midnight ...

In both versions, Jane and Gloria reunite ("Gloria was still up when Jane arrived home about midnight") and then Gloria uses an unheralded pronoun ("Did she poison you again?"). The question of interest is, how accessible was the mutually-known concept cousin at different points in the story? To test this, the reader was interrupted at different test points and asked to quickly decide if "cousin" appeared in the story.

The measure of interest was the difference in recognition time for the common ground concept (cousin) in the concept-absent and concept-present conditions. The waxing and waning of the test concept in the concept-absent condition compared to its accessibility in the concept-present condition indicates the way that readers use common ground to manage the accessibility of information. There were two main results of which I will discuss one, namely, the reunion effect. The reunion of the characters was sometimes sufficient to reduce the accessibility advantage of the concept-present condition. The effect was interpreted as evidence for the role of common ground in text comprehension, suggesting that common ground information between characters becomes available when they unite.

Applying the Subsuming Theory Criterion. This experiment provides evidence that the concept was activated following the reunion. However, it does not provide evidence that it was activated *because* the concept was part of the common ground between the two characters. Therefore, it does not satisfy the subsuming-theory criterion. Mapping this onto the blessed, good paper theory, one can explain the reunion effect with either Theory 1 or 2:

Theory 1: The Subsuming Theory.
(1) Jane returned to the reader's focus of attention.
(2) Information associated with Jane became accessible.

(3) Gloria and Jane were reunited.
(4) The common ground of Gloria and Jane became accessible.

Because cousin is part of the common ground, it became accessible.

Theory 2: The Perfectly-Subsumed Theory.
(1) Jane returned to the reader's focus of attention.
(2) Information associated with Jane became accessible.

Because cousin is highly associated with Jane, it became accessible.

In other words, the reason the concept became accessible might not be the fact that the characters reunited and hence their common ground became more accessible, but the fact that Jane returned to the foreground. Given that Jane is highly related to the concept cousin because the reader knows that she just visited her cousin, this could be a sufficient reason to make cousin more accessible. Indeed, there are several reports of similar effects in the literature, where concepts' accessibility depends on the extent to which they are related to elements in the concurrent focus of the reader (Anderson, Garrod & Sanford, 1983; Glenberg, Meyer & Lindem, 1987; Lesgold, Roth & Curtis, 1979). Because a simpler, subsumed theory can account for the results, it is difficult to draw conclusions that require additional assumptions regarding the role of common ground.

In practical terms, the design confounds two factors: the information associated with the concept before and after reunion with the common ground status of that information. In order to be relevant for the issue of common ground, the experiment should keep the information constant and only manipulate its common ground status. In this example, one could manipulate common ground by having Jane return to the apartment but in one case she reunites with Gloria and in the other she does not (e.g., Gloria just left the apartment). In this way, the information which is common is kept constant (e.g., concepts which are associated with Jane and are part of common ground), but in the reunion case this information is common ground with a present other and in the no reunion case it is not. Alternatively, one could satisfy the subsuming-theory criterion by directly manipulating the common ground status of the information. The experiment would have to demonstrate that given a reunion, a concept (cousin) becomes accessible when it is part of common ground, but it is not accessible when it is not part of the common ground. In the cousin example, this could be achieved by manipulating whether or not Gloria knows that Jane went to dinner with her cousin. The extent to which the concept is associated with Jane would be kept constant, but its common ground status would be manipulated.

It could very well be the case that character reunions indeed activate common ground. Unfortunately, these data do not allow us to conclude this because they do not rule out subsumed theories which do not assume common ground. Consistent

with my argument, McKoon, Gerrig and Greene (1996) report very similar character reunion effects, but explain them in terms of strength association between characters and concepts (i.e., the subsumed theory), not in terms of common ground (i.e., the subsuming theory). So does the reunion of characters make their common ground more accessible? The conclusion from this analysis should be that we don't know.

CAN COMMON GROUND BE UNCONFOUNDED?

One might argue that it is impossible to unconfound common ground experimentally precisely because it is inherently confounded. If this is true, then we need not bother conducting experiments to test the theory of common ground because the conclusions would not bear on the issue. I propose that it is possible to unconfound common ground and therefore it is possible to test theories which have it as a central concept. In order to do this consistently, we need to apply the subsuming theory criterion. Practically speaking, this means that we must keep common information constant while varying its common ground status. I will apply the criterion by briefly reviewing a study which seems to unconfound common ground and explain why it actually does not. Then I will mention studies which satisfy the criterion as positive examples of the unconfounding of common ground.

Do We Use Common Ground to Understand Demonstrative Reference?

Clark, Schreuder and Buttrick (1983) investigated the role of common ground in the comprehension of demonstrative reference. In the first experiment the experimenter presented subjects with a picture of flowers and asked, "How would you describe the color of this flower?" Given that the picture had flowers of different colors, the reference "this flower" was ambiguous. The main goal of the study was to demonstrate that addressees will accept such a potentially ambiguous reference. If they don't, then addressees should not be able to interpret the expression "this flower" when the picture has different flowers. To test this, the experiment contrasted the responses to two different pictures. In one picture the different flowers were equally salient and in the other picture one type of flower was much more salient than the other. The results were straightforward: With the picture that had a highly salient flower, most of the subjects immediately chose that flower as the referent and provided its color as a response. In contrast, subjects rarely responded with that particular color when the flowers did not vary in salience, and instead they tended to ask for clarification. The paper described two theories which can account for the results. A common ground account can explain

the differential results by assuming that subjects selected the flower which was mutually salient to them and to the questioner. A "perceptual" theory would suggest that subjects identified the salient flower as the referent simply because it was salient to them. While both the perceptual and the common ground theories could account for the first two experiments in the paper, the last two experiments attempted to uniquely support the common ground theory. Though they appear to do this, these experiments do not unconfound common ground.

Consider Experiment 4. The experimenter presented people with a picture of then President Reagan and his director of the Office of Management and Budget, Stockman. He then asked either, "You know who this man is, don't you?" or, "Do you have any idea at all who this man is?" People were much more likely to identify Reagan as "this man" after the first question, and Stockman after the second question. This result clearly rules out a purely perceptual account, because had people identified "this man" on the basis of salience, they should have selected the same person with both questions. This experiment looks like it directly supports the claim that comprehenders use common ground, in this case common ground by virtue of community membership. The very use of presuppositions in the questions seems to make use of mutuality of knowledge. In the one case, the speaker's question can be paraphrased to mean: "Indicate the man which I have a good reason to believe that you are familiar with as a member of our community," whereas in the other case the question can be paraphrased as: "Indicate the man which I have a good reason to believe that you are unfamiliar with as a member of our community." The differential results for the two questions, then, might reflect the use of common ground.

It might very well be the case that when addressees comprehend questions with such presuppositions they do use common ground. Unfortunately, the experiment's design does not allow such a conclusion because it does not satisfy the subsuming theory criterion. In a nutshell, the source of the variation could be the fact that the addressees were more familiar with Reagan than with Stockman (i.e., the information), not the fact that the speaker has a good reason to believe that the addressee is more familiar with Reagan (i.e., the fact that the information is common ground). In practical terms, in order for the experiment to supply relevant evidence it must vary only the common ground status of the information. In this case one would want to vary the community membership of the speaker and addressee – comparing a case when they do belong to the same community to a case when they don't.

Another way to look at this experiment is in terms of linguistic co-presence, which is more directly related to the way the experiment is framed. When the speaker utters the question, the presupposition becomes linguistically co-present for the speaker and the addressee. Perhaps the fact that two different presuppositions were co-present lead addressees to two different answers.

Although possible, this conclusion is not warranted by the data. Again, the reason is that this experiment does not directly contrast co-presence with mere presence—a co-present presupposition is also present for the addressee and this might have been sufficient for the results. It might turn out that such a requirement is too difficult to fulfill with the use of presuppositions as the linguistic instrument because it might result in awkward situations. If this is the case, then presuppositions might not be an adequate tool to investigate common ground.

One might misinterpret my argument as concluding that language users do not use common ground. This would be misguided. Instead, my claim is that experiments which do not distinguish between the subsuming and the subsumed theory cannot provide data which are relevant for the issue of common ground. Consequently, the answer to the question "Did subjects in the demonstrative reference experiments use their common ground with the speaker?" should be that we don't know, just as we don't know if common ground played a role in the unheralded pronoun experiments described earlier. Therefore, these experiments should not alter our beliefs with respect to the issue of common ground.

Unconfounded Examples

It might be useful to contrast experiments which do not satisfy the criterion with those that do. In practical terms, the design of all unconfounded experiments kept the relevant information constant while varying only whether or not the information is common ground. Consider Gerrig and Littman's (1990) study on how readers use community membership to disambiguate utterances. They presented Yale students with scenarios such as the story about a football game between Yale and Brown. The readers had to imagine that after the game either a friend or a stranger says to them: "That game was a disaster." In Experiment 1 participants were more confident in their interpretation of the speaker's utterance when the speaker was a friend than when he or she was a stranger. Further experiments demonstrated parallel results with comprehension difficulty measures. The results were interpreted in terms of use of common ground by virtue of community membership. This experiment satisfies the subsuming theory criterion because it manipulates the community membership and contrasting interpretations for speakers which are members of the same community with speakers who are not.

While Gerrig and Littman investigated comprehension, production studies have been more likely to satisfy the subsuming theory criterion. Consider the way Fussell and Krauss (1992) used gender as a manipulation of community membership. They first established that the population of their participants assume that certain categories are more familiar to men (e.g., tools), and others more familiar

to women (e.g., kitchen utensils). Then they had participants describe objects from these categories for either a male or a female. They hypothesized that if speakers use common ground to tailor their message to their addressee, they should provide more detailed identifying information when the addressee might be less familiar with an object than when the addressee's gender category suggests more familiarity with the objects. Similar to Fussell and Krauss' design, Isaacs and Clark (1987) compared the way New Yorkers described New York City landmarks depending on whether or not their addressee was another New Yorker. Such a design can address the issue of common ground because it satisfies the subsuming theory criterion. It keeps constant the speaker's knowledge of certain objects and only varies whether or not this knowledge is shared with the addressee by manipulating the addressee's membership in the relevant community.

Horton and Keysar (1996) applied the criterion to address common ground by virtue of perceptual co-presence, and demonstrated how perceptual presence can be teased apart experimentally from perceptual co-presence in the context of speakers' use of common ground. We did this by comparing the extent to which speakers use context when it is shared with their addressees to their use of the same context when it is privileged to them. This method allowed us to distinguish between production processes which are sensitive to common ground and processes which are not. By using the proposed criterion, then, one can directly test the theory of common ground.

CONCLUSIONS AND IMPLICATIONS

One should not conclude from this study that common ground plays no role in language use, or that language users rely only on "self" knowledge. The argument is only about the extent to which a particular design has the potential to provide evidence relevant for theories that assume common ground, and the proposal is with respect to the way we can tighten rules of evidence to conform to the general principle of parsimony. One might also think that I am suggesting that experiments on common ground should rule out alternative explanations in general. This interpretation of my argument would be misguided. The possibility of alternative explanations is not the problem here. Subsumed alternative explanations are, because they contribute a systematic source of confusion. Experiments need to demonstrate the force of the subsuming explanation, not the subsumed one.

The Importance of the Proposed Criterion. The reason that a subsumed alternative explanation is a potential problem has to do with the principle of parsimony: Whenever a subsumed explanation is possible, it should take precedence over the explanation which includes it. The subsumed explanation, by definition,

makes fewer assumptions. This is why the subsuming-theory criterion is a necessary condition. It assumes that parsimony takes precedence for blessed, good paper theories, for common ground-based theories, and for any attempt to link experimental results with theory. If one refuses to accept the subsuming-theory criterion as a necessary condition then one must also refuse the analogous condition for the blessed, good paper theory of publication. One would be forced to conclude that good papers are accepted for publication because of the saint's blessings. People who reject the subsuming, saint-driven theory when no evidence is provided that Joseph's paper was accepted because of the saint's blessings should not be willing to accept the conclusion that information was used because it was mutually shared when the experiment does not satisfy the subsuming-theory criterion.

One might misinterpret the argument to be that we should have more parsimonious theories in pragmatics, or that more parsimonious theories are more likely to be correct than less parsimonious theories. I am not advocating parsimonious theories. I am arguing for the use of parsimony when an experiment does not distinguish between theories. In that case, the more parsimonious theory should be preferred.

The Benefits from Applying the Criterion. The implications of the proposed criterion go beyond the concerns of the specific area of psycholinguistics. The concepts of common ground and mutual knowledge are central not only to the cognitive study of language but also to other related areas such as the social psychological study of interaction (Fussell & Krauss, 1989; Krauss, 1987; Krauss & Fussell, 1991; Slugoski, Lalljee, Lamb, & Ginsburg, 1993) and the study of language development (Ackerman, Szymanski, & Silver, 1990; Kail & Hickmann, 1992; Newman, 1986). Given the diversity of the use of this concept, a unified standard of evidence might be useful.

Another benefit from using the subsuming theory criterion is that it might allow us to distinguish between real and apparent cases of the use of common ground. Brown and Dell (1987) used a logic analogous to that presented here and were able to conclude that what seemed to be a communicative-based effect was not. They demonstrated that what looked like a speaker's attempt at being informative (Grice, 1975) was only apparent. Subjects' behavior in their experiment seemed to be motivated by an attempt to be in line with Grice's informativeness maxim. To test this, Brown and Dell compared a case when the addressee could benefit from the information with a case where the addressee was already informed. Speakers behaved the same way, regardless of whether or not the addressees were informed. Brown and Dell's design, then, allowed them to demonstrate that an effect that could have been attributed to a communicative behavior was actually independent of it. In addition to distinguishing between the actual

and apparent role of common ground, one could use the proposed criterion to distinguish between aspects of a process where common ground does play a role and aspects where it does not (e.g., Horton & Keysar, 1996; Keysar & Pack, 1993).

In general, the use of the proposed criterion will benefit the field of pragmatics by helping to re-evaluate our beliefs about the role common ground plays in language use. It will allow a new integration of past research, and it should be useful in the design of future experiments. One does not need experiments to convince people that interlocutors use common ground when they communicate. This intuitive assumption makes a lot of sense from the outset, and in one form or another it is probably correct. Yet experiments can provide evidence which is relevant for this intuition only if their design satisfies the subsuming theory criterion.

Acknowledgement: An earlier version of this paper was presented at the 36th annual meeting of the Psychonomic Society in Los Angeles (Keysar, 1995). The preparation of this paper was supported by PHS, grant R29 MH49685 to the University of Chicago, Boaz Keysar, Principal Investigator. I am grateful to PHS for this support. I am indebted to my graduate students and research assistants for helping me to think about these issues, to Kate Baldwin, Dale Barr, Linda Ginzel, Sid Horton, Matt McGlone, Howard Nusbaum, Adam Schorr, Joseph Stern and Tom Trabasso for insightful comments on earlier drafts and Jennifer Balin for her invaluable assistance. I am grateful to Steve Greene for sending me the complete list of items used in Greene, Gerrig, McKoon and Ratcliff (1994) and especially for his cooperative approach throughout.

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.

Anderson, A., Garrod, S.C., & Sanford, A.J. (1983). The accessibility of pronominal antecedents as a function of episodic shifts in narrative text. *Quarterly Journal of Experimental Psychology, 35A*, 427–440.

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

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

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

Clark, H.H. & Brennan, S.E. (1991). Grounding in communication. In L. B. Resnick, J. Levine, & S. D. Teasley (Eds.), *Perspectives on socially shared cognition* (pp. 127–149). Washington, D.C.: APA.

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* (pp. 57–63). 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.

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.

Fussell, S.R. & Krauss, R.M. (1992). Coordination of knowledge in communication: Effect of speakers' assumptions about what others know. *Journal of Personality and Social Psychology, 62*, 378–391.

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

Gibbs, R.W. (1987). Mutual knowledge and the psychology of conversational inference. *Journal of Pragmatics, 11*, 561–588.

Glenberg, A.M., Meyer, M., & Lindem, K. (1987). Mental models contribute to foregrounding during text comprehension. *Journal of Memory and Language, 26*, 69–83.

Goodman, N. (1979). *Fact, fiction and forecast*. Cambridge, MA: Harvard University Press.

Graesser, A.C., Bowers, C., Bayen, U.J., & Hu, X. (in press). Who said what? Who knows what? Tracking speakers and knowledge in narrative. In W. van Peer, E. Andriga, D. Schram, & E. Tan (Eds.), *Narrative perspective: Cognition and emotion*.

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*, Vol. 3: *Speech acts* (pp. 41–58). New York: Academic Press.

Harman, G. (1977). Review of *Linguistic Behavior* by Jonathan Bennett. *Language, 53*, 417–424.

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

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

Johnson-Laird, P.N. (1982). Mutual ignorance: Comments on Clark and Carlson's paper. In N. V. Smith (Ed.), *Mutual knowledge* (pp. 40–45). London: Academic Press.

Kail, M. & Hickmann, M. (1992). French children's ability to introduce referents in narratives as a function of mutual knowledge. *First Language, 12*, 73–94.

Keysar, B. (1995). *In defense of common ground.* Paper presented at The 36th Annual Meeting of the Psychonomic Society, Los Angeles, November.

Keysar, B. & Paek, T.S. (1993). *Definite reference and mutual ignorance*. Paper presented in The 34th Annual Meeting of the Psychonomic Society. Washington, D.C.

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.

Lesgold, A.M., Roth, S.F. & Curtis, M.E. (1979). Foregrounding effects in discourse comprehension. *Journal of Verbal Learning and Verbal Behavior, 18*, 291–308.

McKoon, G., Gerrig, R.J., & Green, S.B. (1996). Pronoun resolution without pronouns: Some consequences of memory-based text processing. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 22*, 919–932.

Newman, D. (1986). The role of mutual knowledge in the development of perspective taking. *Developmental Review, 6*, 122–145.

Schiffer, S.R. (1972). *Meaning*. Oxford: Clarendon Press.

Slugoski, B.R., Lalljee, M., Lamb, R., & Ginsburg, G.P. (1993). Attribution in conversational context: Effect of mutual knowledge on explanation giving. *European Journal of Social Psychology, 23*, 219–238.

Sperber, D. & Wilson, D. (1982). Mutual knowledge and relevance in theories of comprehension. In N. Smith (Ed.), *Mutual knowledge* (pp. 61–87). London: Academic Press.

Stalnaker, R. C. (1978). Assertion. In P. Cole (Ed.), *Syntax and semantics*, Vol. 9: *Pragmatics*. New York: Academic Press.