SPEAKERS’ OVERESTIMATION OF THEIR EFFECTIVENESS
Boaz Keysar and Anne S. Henly
The University of Chicago

Abstract—Successful communication depends in part on an ability to anticipate miscommunication. We investigated speakers’ ability to gauge their addressees’ understanding. Participants in our experiments were asked to say ambiguous sentences while attempting to convey a specific intention to their addressee. When they estimated the addressee’s understanding of the intended meaning, they showed a consistent tendency to overestimate their effectiveness. They expected the addressee to understand more often than the addressee actually did. In contrast, overhearers who were informed about the speakers’ intention did not systematically overestimate the speakers’ effectiveness. Our findings suggest that when speakers monitor their own utterances, they do not act as unbiased observers. Instead, they underestimate the ambiguity of their own utterances and overestimate the extent to which their disambiguating cues make their intention transparent. Such overestimation could be a systematic source of miscommunication in natural conversation, and should be accounted for by any theory of language production.

An American psycholinguist (M.G.) participated in a conference in Canada. At the banquet, her table wanted to order wine that would match the food. She approached the waiter and said, “Excuse me, but could you tell me what the entree is?” He looked a bit baffled at first and then very slowly said, “Well, it’s the course that comes after the salad but before the dessert.” Apparently M.G. did not appreciate the ambiguity of her own question. In the study we report here, we investigated the possibility that speakers in general underestimate the ambiguity of their utterances, and that consequently they overestimate their ability to convey what they intend. We suggest that the very attempt to communicate contains the seeds of miscommunication.

MONITORING AND AMBIGUITY

Speakers can disambiguate their utterances by using context, prosody, clarification, and other cues. Prosody is an especially powerful tool for indicating pragmatic intent. For example, a slight stress on “love” in “I love you” could be an appropriate statement on Valentine’s Day, but the same words with a rising, questioning intonation on the “you” could ruin the occasion. However, some ambiguities are relatively difficult to resolve. For example, the sentence “The daughter of the man and the woman arrived” could either mean that the daughter and the woman arrived or that only the daughter arrived. Speakers can prevent such syntactic ambiguities, in principle, by varying prosody—an exaggerated pause after “man” to convey one meaning, a pause after “woman” to convey the other (e.g., Lehiste, 1973; Lehiste, Olive, & Streeter, 1976). Indeed, such prosodic cues can be used to disambiguate many utterances (e.g., Allbritton, McKoon, & Ratcliff, 1996; Price, Ostendorf, Shattuck-Hufnagel, & Fong, 1991).

But do speakers spontaneously disambiguate their utterances? Allbritton et al. (1996) found that they do not. They asked professional speakers and untrained college students to read syntactically ambiguous sentences such as “When you learn gradually you worry more,” and convey either that you gradually learn or that you gradually worry. Neither group of speakers produced sufficiently distinct prosody to allow other individuals to distinguish between the two meanings, except that the professional speakers did so when they were explicitly instructed to disambiguate the utterance. Thus, even though speakers may in principle use prosodic cues to clarify their intention, they do not do so spontaneously.

Yet the fact that speakers produce ambiguous utterances does not necessarily mean they will miscommunicate. Speakers might allow some ambiguity in their speech, but monitor the communication in order to anticipate a misunderstanding and clarify their intention. In this case, their success will depend on their ability to predict or detect misunderstanding. The question then becomes, to what extent are speakers able to gauge the understanding of their addressees?

Speakers monitor their utterances to detect potential errors (De Smedt & Kempen, 1987; Laver, 1980; Levelt, 1983), such as errors in register, word choice, social acceptability, and the match between their intention and what they actually say. When they detect an error, they initiate either an internal revision or an overt repair (Berg, 1992; Levelt, 1989). In the present study, our focus is on speakers’ monitoring of what they say and their ability to detect a listener’s misunderstanding. When speakers gauge addressees’ understanding, there are three possible outcomes:

1. If they are well calibrated, they know when their addressees have understood them and when their addressees have not understood them.
2. If they overestimate their effectiveness, they think they are understood when they are not.
3. If they underestimate their effectiveness, they think they are not understood when in fact they are.

Speakers cannot always be perfectly calibrated. Random error could cause them to overestimate and underestimate occasionally. If such error is indeed random, one would not expect a difference in the rates of over- and underestimation. But if speakers overestimate more often than they underestimate their effectiveness, this would suggest a systematic, not random, error.

In Experiment 1, we evaluated speakers’ ability to accurately assess their utterances’ effectiveness. Each speaker attempted to convey a particular meaning of an ambiguous sentence to an addressee and then judged what the addressee understood. By comparing the actual understanding of the addressees with the speakers’ expectations, we could measure speakers’ over- and underestimation.

Speakers in Experiment 1 indeed tended to overestimate their effectiveness, so the goal of Experiment 2 was to compare two possible reasons for such a tendency. One possibility is that knowing the intended meaning of an utterance makes that meaning more accessible than other interpretations, leading the speaker to believe the addressee will also...
Speakers’ Overestimation

find it a more likely interpretation of the utterance. A second possibility is that the very attempt to convey a particular intention, via their manner of speaking, leads speakers to overestimate their effectiveness. For example, the cognitive resources speaking requires might not leave sufficient capacity available to adequately monitor for ambiguities. To evaluate these two possibilities, we used “overhearers” in Experiment 2. These participants received exactly the same information as the speaker, but rather than produce the utterances themselves, they listened to recordings of the speaker’s utterances and then judged the addressee’s understanding. If knowledge of the intended meaning is the reason for speakers’ overestimation, then such overhearers should show the same bias because they also know the intention. If, however, the very act of speaking is responsible for the effect, then overhearers should not show a tendency to overestimate. We report the two experiments together to allow easy comparison of the results for speakers and overhearers.

METHOD

Participants

The participants were native English speakers, none of whom had any history of language disorder nor any speech or hearing impairment. Seven speaker-listener pairs were replaced for not following instructions. Eighty people participated as speaker-listener pairs in Experiment 1, and 37 additional individuals participated as yoked overhearers in Experiment 2 (the recordings of 3 speakers were not usable because of equipment failure).

Materials

Twelve syntactically ambiguous sentences (e.g., “Angela shot the man with the gun”) and four lexically ambiguous sentences (e.g., “Currency is no longer exchanged by the banks”) served as experimental items. Each sentence was preceded by a short, disambiguating scenario. The scenario ended with a question to which the experimental sentence was the answer. For example, “Angela killed the man with the gun” appeared after both scenarios in Table 1. The first scenario described a man with a briefcase and another with a gun, and ended with “Which man did Angela kill?” In the second scenario, Angela selected a gun to kill the man, and it ended with “Which weapon did Angela kill the man with?”

The two meanings of each sentence varied in their relative “strength,” or dominance, when presented alone. For example, in “Currency is no longer exchanged by the banks,” the institution meaning of banks comes to mind more easily than the riverside meaning. We established the relative bias for each meaning in a separate normalizing study in which 16 additional participants read each experimental sentence without the context. They indicated which of the two meanings came to mind first when they read the sentence. The proportion of participants to select each meaning served as a measure of its relative dominance. Sentences varied in the differential dominance of the two meanings; some items had a highly dominant meaning, some a moderately dominant meaning, and some had no dominant meaning.

We collected the scenarios in booklets that we gave to the speakers. Each booklet presented only one meaning of each sentence, resulting in two different booklet types with 16 experimental items per booklet. Each scenario and corresponding sentence appeared on a separate page. The booklets were equated for an average dominance bias of 50% across items.

Design and Procedure

The members of each pair of participants were randomly assigned to the speaker and listener roles. They each received a booklet that began with instructions and an example of an ambiguous sentence, “Rick moved the grill under the porch.” The instructions explained the ambiguity and illustrated it by providing the two possible meanings, “The grill under the porch was moved by Rick” and “The grill was moved under the porch by Rick.” The speaker’s booklet provided him or her with the test sentences and their corresponding scenarios; the listener’s booklet included only the item numbers and blank spaces to fill in the meaning of each item (choice of A or B). The two different meaning paraphrases for each item were provided to the participants only after the speaker said the sentence.

The speaker and listener sat opposite each other. On each trial, the speaker first silently read the disambiguating scenario and question while the target sentence was kept covered. Then the speaker uncovered the sentence and said it while facing the listener. The speaker knew that the listener did not have the disambiguating scenarios, and that the goal was to convey the contextually appropriate meanings to the listener. The listener was informed that the speaker was equally likely to intend either possible meaning of each sentence.

As soon as the speaker finished saying each sentence, he or she began the task of assessing the listener’s understanding. The speaker was instructed to turn around immediately after saying the utterance, and so could not use visual clues such as indications of puzzlement by the listener. The experimenter then projected the two possible meanings on the wall, and the speaker indicated which he or she thought the listener had understood. At the same time, the listener selected which meaning he or she believed the speaker actually intended, and indicated his or her certainty on a 5-point scale (1 = very uncertain and 5 = very certain).

The speakers in Experiment 1 were recorded on digital audiotape for use in Experiment 2. New participants played the role of overhearers in Experiment 2. Each was yoked to an individual speaker-listener pair, and was matched to the speaker’s gender. Each overhearer received the same booklet the speaker had, as well as an overview of the original experiment in order to understand the context for the task. On each trial, the overhearer first read the scenario and question, and then uncovered the corresponding sentence just as the speaker had in Experiment 1. Then the experimenter played the speaker’s recording and projected the two meanings on the wall. The overhearer performed the same task as the speaker had, indicating which meaning he or she thought the original listener had understood.

RESULTS AND DISCUSSION

Listeners’ Accuracy

Overall, listeners were not very accurate. Their mean accuracy was 61%, which was statistically different from chance (50%), t(39) = 6.35, p < .001. They were more accurate with syntactic than with lex-

1. We initially reasoned that speakers would recognize the impossibility of disambiguating lexical meaning with prosody and would therefore not overestimate addressees’ understanding of lexically ambiguous items. As it turns out, we underestimated speakers’ ability to overestimate, as they showed the bias even with lexically ambiguous sentences.
Speakers’ Over- and Underestimation

Overall, speakers believed that their addressees understood their intended meaning in 72% of the cases (76% and 59% for syntactically and lexically ambiguous sentences, respectively). The difference between their predictions and listeners’ actual accuracy (61%) was significant, t(39) = 4.36, p < .001. Figure 1 presents speakers’ predictions of their addressees’ identification accuracy, contingent on the addressees’ actual accuracy. The diagonal represents a calibrated correspondence between speakers’ expectations and addressees’ performance. Most speakers’ predictions fall above this line, showing an overestimation bias.

We define overestimation as the probability of predicting that the addressee understood given that the addressee did understand. We define underestimation as the probability of predicting that the addressee did not understand when the addressee did understand. Overall, speakers overestimated more than they underestimated (Fig. 2). When the addressees did not understand the intention (n = 249), the speakers thought they did in 46% of the cases. In contrast, when the addressees did understand the intention (n = 387), the speakers said that did not in only 12% of the cases. Eighty percent of the speakers showed a tendency to systematically overestimate, 10% of the speakers showed a tendency to systematically underestimate, and 10% showed no systematic bias. Across speakers, the overestimation measure was significantly larger than the underestimation measure, t(39) = 6.74, p < .001.

We considered the possibility that overestimation was caused by low-dominance items. Listeners might have defaulted on dominant meanings, in which case they would have tended to be wrong when less dominant meanings were intended and to be right when more dominant meanings were intended. In this case, it would have been

---

2. We also conducted a signal detection analysis on these data. The conclusions were the same as we report here, but we opted for the current analysis because it is more transparent and because it is not clear that all assumptions hold for a bona fide signal detection analysis.
Speakers’ Overestimation

easier for speakers to overestimate their success with lower dominant than with higher dominant meanings. The results indicate otherwise; speakers’ bias correlated somewhat with dominance ($r = .5$, $p < .01$), but with a stronger tendency to overestimate with more dominant than with less dominant meanings.

**Overhearers Versus Speakers**

In contrast to speakers, overhearers showed no tendency to overestimate. Overall, they expected addressees to understand the intended meaning in only 56% of the cases. As Figure 2 illustrates, although speakers tended to overestimate more than underestimate, by a difference of 34 percentage points, there was only a 2-percentage-point difference between over- and underestimation for overhearers (30% and 28%, respectively), $t(36) = 0.016$, $p = .7$.

Overall, speakers overestimated more when they produced syntactically ambiguous sentences ($M = 55\%$) than when they produced lexically ambiguous ones ($M = 29\%$). This suggests that speakers appreciated the impossibility of disambiguating a lexically ambiguous sentence, and thus were less likely to overestimate their ability to convey its intended meaning. Still, the pattern of overestimation bias was the same for the two types of items. In both cases, speakers tended to overestimate more than underestimate, with a mean difference of 42 percentage points for syntactically ambiguous items, $t(39) = 6.77$, $p < .001$, and a 22-percentage-point difference for lexically ambiguous items, $t(37) = 3.69$, $p < .001$. In contrast to speakers, overhearers did not overestimate much more with syntactically ambiguous items than with lexically ambiguous items ($Ms = 33\%$ and 24%, respectively). They also did not show an overestimation bias with either item type. The mean overestimation bias was 2 percentage points for syntactically ambiguous items, $t(36) = 0.32$, $p = .7$, and 7 percentage points for lexically ambiguous items, $t(31) = 1.24$, $p = .2$. The two types of ambiguities, then, showed the same pattern for overhearers.

**GENERAL DISCUSSION**

Our results are clear. Speakers tended to overestimate their effectiveness and expected their addressees to understand their intentions more than was warranted. This was the case even though the speakers were fully aware of the ambiguity of the expressions at the time they

![Fig. 2. Average probability of overestimation and underestimation by speakers in Experiment 1 and by overhearers in Experiment 2. The probability of overestimation was calculated as the probability of saying the listener understood when he or she did not understand; the probability of underestimation was calculated as the probability of saying that the listener did not understand when he or she did.](image-url)
assessed their addressees’ understanding. Experiment 2 shows that the overestimation bias is unique to speakers: Overhearers who possessed the same contextual information as the speakers and heard what the speakers said did not tend to systematically overestimate the effectiveness of the speakers.

Our results are relevant to theories of speech production because they are informative regarding the way speakers monitor their own speech. We have shown that if indeed speakers monitor by serving as their own listeners (Levett, 1983), the result of such monitoring is systematically different from the understanding of a true listener. Speakers perceive their own utterances as more transparent than they really are.

Did our speakers simply engage in impression management? It is possible that they realized they failed to communicate, but said they were successful to maintain the impression that they were effective. If this is the case, then our results might be due to the specific demands of the experiment. We suggest that there are at least two reasons to discount such a possibility. First, nearly a third of the time (28%), our speakers did report failure to communicate, which is a fairly high rate if they were attempting to present an “effective communicator” self. Second, if our speakers falsely reported that they communicated effectively, then one would expect such a tendency to be strongly attenuated with increased public accountability. Accountability typically induces the opposite strategy, as it promotes self-protective presentation strategies, such as making safer, more secure decisions (Schlenker, 1980, 1987), and causes actors to lower public expectations (Maracek & Mettee, 1972). Therefore, if our speakers really believed that they were not effective, then we would expect their predictions to become more calibrated under conditions of increased accountability. In a related study, we asked speakers to predict how effective they would be at identifying their own intention from a recording of their utterance (Keysar & Henly, 1998). These subjects knew that they would be publicly accountable for their predictive accuracy, because they would have to attempt to identify their own intended meaning the next day. Despite this increased accountability, they overestimated their ability to identify their own intention to the same degree as they overestimated listeners’ understanding. This suggests that our speakers believed they were effective and were not strategically engaged in impression management.

Speakers’ overestimation might share an underlying mechanism with the phenomenon of the illusion of transparency (Gilovich, Savitsky, & Medvec, 1998; Keysar, 1994, 2000; Keysar, Ginzel, & Bazerman, 1995). In a series of experiments (Keysar, 1994), subjects read about a protagonist who hated a restaurant and left a friend a sarcastic note saying, “That restaurant was marvelous, just marvelous.” Participants believed that the protagonist’s friend would perceive his sarcastic intention even when the addressee lacked the information about the restaurant, which was crucial to disambiguate the note. They behaved as if the protagonist’s intention was transparent. Similarly, Gilovich and his colleagues showed that people tend to overestimate the extent to which their internal states are transparent to others. Subjects in their experiments thought that their private preferences were clearer to observers than they really were. They behaved as if their internal sensations, emotions, and thoughts “leaked out” in public. In much the same way, speakers in our experiments took their own intentions as relatively transparent, and underestimated the ambiguity of their utterances.

Speakers’ tendency to overestimate their effectiveness could also be related to the illusion of control (e.g., Langer, 1975; Thompson, 1999; Thompson, Armstrong, & Thomas, 1998). Several studies have demonstrated that actors overestimate the extent to which they can control chance outcomes via their actions. For example, people prefer to use lottery tickets for which they have selected the numbers even when they have the option to switch to tickets with better odds of winning. They behave as if their active selection of the numbers exerts some control over the outcome (Langer, 1975). Interestingly, an illusion of control is unique to actors. Observers do not think that actors are more effective than they really are, just as our overhearers did not overestimate the effectiveness of the speakers.

The cognitive demands of speaking might also explain the difference between our speakers and overhearers. Perhaps the act of speaking taxes the cognitive system, leaving relatively few resources for monitoring one’s own utterances. Consequently, speakers might have difficulty modeling the perception of their utterance from the listener’s perspective. Such a process would lead to the systematic bias we observed because it should increase speakers’ tendency to believe that the addressee understands the utterance when it is relatively ambiguous.

The difference between speakers and overhearers might also be due to speakers’ privileged knowledge of the prosodic cues they employ to disambiguate their utterances. Speakers are privy to how they intend their prosody to indicate their intended meaning. For example, when speakers say, “Angela killed the man with the gun,” they might pause after “man” in order to convey that she used the gun. Even when the pause is not distinctive enough, the speakers might still perceive it as sufficient simply because they know how they intended it. In contrast, overhearers do not know which disambiguating tools the speakers have chosen; they can only hear the utterance itself. The lack of that knowledge might spare overhearers from a systematic overestimation.

To the extent that our findings reflect speakers’ tendency to take their own utterances as relatively transparent, these findings suggest that at least part of miscommunication could be systematic. If speakers believe that their addressees understand them, they might be less likely to verify that the addressees indeed arrive at the intended meaning. If speakers receive clear feedback that they have disambiguated, they should be able to revise their utterances (e.g., Traxler & Gernsbacher, 1992, 1993), but it is not clear to what extent addressees tend to provide such clear feedback. For instance, our listeners’ confidence was unrelated to their accuracy; they were just as confident when they did not understand the speaker’s intention as when they did. This suggests that feedback to the speaker might have little diagnostic value, and might provide the speaker with few reliable opportunities to learn how ambiguous his or her utterances are.

The potential for miscommunication might be increased by addressees’ tendency to interpret utterances egocentrically (Keysar, Barr, & Horton, 1998). Recent studies indicate that addressees do not restrict their interpretation to information they know to be mutually shared with the speaker (Keysar, Barr, Balin, & Brauner, 2000; Keysar, Barr, Balin, & Pack, 1998). When speakers referred to objects, addressees often identified the referents from their own egocentric perspective, sometimes even identifying objects that were visible only to them and thus could not have been the intended referents. Such egocentric processing by addressees, together with speakers’ tendency to take their utterances as clear and relatively unambiguous, could lead to systematic failure of communication. Moreover, our experiments may underestimate the extent of speakers’ overestimation bias. Speakers in our experiments were highly sensitized to the potential ambiguity of
Speakers’ Overestimation

their utterances. They knew that the sentences were ambiguous, they could anticipate the type of ambiguity, and they were provided with the two potential meanings of each sentence. This should have greatly facilitated their ability to accurately gauge their addressees’ understanding. In contrast, in natural conversation speakers might often be completely unaware of alternative interpretations for their utterances. Consequently, they may be more likely than our participants to believe that their addressees understand them when they have not.

A potential drawback of our study is that the speakers were provided with sentences to read—they did not decide on the actual wording, only on the way to utter the words. Perhaps when speakers generate their own sentences they avoid the kind of ambiguity our sentences presented. If speakers naturally produce unambiguous utterances, then listeners would understand them and speakers would not have occasion to overestimate their effectiveness in natural conversation. Although it is possible that speakers do not on occasion produce syntactic, lexical, and pragmatic ambiguities, this is highly unlikely. Indeed, Ferreira and Dell (2000) showed that speakers’ choice of syntactic structures is not sensitive to the potential ambiguity that the sentence presents to the listener. No one knows the prevalence of ambiguity in natural conversation, but we suggest that speakers are just as likely to overestimate their effectiveness when they produce ambiguous sentences spontaneously as when they read ambiguous sentences out loud. In fact, the overestimation effect might be even larger with self-generated sentences because of the additional demand that the generation of speech puts on cognitive resources.

We propose that, in general, speakers tend to overestimate their effectiveness. This could be a systematic source of miscommunication that may go undetected by the speakers themselves. The waiter’s effectiveness. This could be a systematic source of miscommunication that the generation of speech puts on cognitive resources.

REFERENCES


Acknowledgments—This project was supported by Public Health Service Grant R29 MH49685 to the University of Chicago. We would like to thank Anetta Jedraszko and Erik Schechter for their help with the experiments, and Susan Goldin-Meadow, Josh Klayman, Howard Nusbaum, Sue Paik, and Mary Yoon for helpful feedback.