8. Self-Anchoring in Conversation: Why Language Users Do Not Do What They "Should"

Boaz Keysar and Dale J. Barr

An old story about J. Edgar Hoover illustrates the miscommunication of intentions. When Hoover noticed that the text of a memo his secretary typed had spilled into the page margins, he scribbled a note to her: "Watch the borders." The next day agents were on high alert on the border with Mexico. The question we'd like to address in this paper is this: Why did Hoover miscommunicate?

We propose that much miscommunication is systematic. It results from the nature of language use. We attempt to demonstrate our point by noting the similarities between the way people make decisions and the way they use language. As in decision making, uncertainty is inherent in the way people use language. The way language users overcome this uncertainty leads to systematic errors.

When we attempt to understand what speakers mean, we must infer what they mean from what they say. This is because all utterances are ambiguous. "Borders" is lexically ambiguous, but ambiguity in language goes beyond lexical ambiguity. In fact, everything people say is ambiguous because it can convey more than one intention. To overcome this inherent ambiguity, we propose that language users rely on certain heuristics of language use. As with other heuristics, they are generally successful but they occasionally lead to systematic error.

In this chapter, we propose that speakers, addressees, and overhearers reduce the uncertainty of linguistic utterances by using an anchoring and adjustment heuristic. We review evidence that language users tend to anchor on their own perspective and attempt to adjust to the perspective of others. These adjustments are typically insufficient, and can occasionally cause miscommunication.

COOPERATION AND PERSPECTIVE TAKING

Research on language use makes strong normative assumptions. It assumes that people adhere to certain normative rules of conversation to avoid misunderstanding. Grice (1975) argued that language users follow these rules and assume that others do too because conversation is governed by a "cooperative

The research described here was supported by PHS, grant R29 MH49685 to Boaz Keysar, and an NSF graduate fellowship to Dale Barr. We thank Linda Ginzel and Carmen Marti for their comments.

https://doi.org/10.1017/CBO9780511808098.010 Published online by Cambridge University Press

principle." Clark and colleagues (e.g., Clark, 1996; Clark & Carlson, 1981; Clark & Marshall, 1981; Greene, Gerrig, McKoon & Ratcliff, 1994) argue that for conversation to be truly cooperative, the production and comprehension of linguistic utterances should be governed by mutual knowledge – by information that interlocutors know they share. Thus, cooperative speakers should design utterances bearing in mind the other's thoughts, beliefs, assumptions, and so on. This is known as the *principle of optimal design*. Addressees, in turn, should assume that speakers' utterances are optimally designed for them in this way. In short, the cooperative interlocutor should take into account the other's perspective.

For example, when your annoyed friend says to you, "Nice job," after you wrecked her car, you interpret her utterance with respect to information that is mutually known: You both know that you wrecked her car, that a car is an expensive item, that people are rarely happy when they lose their means of transportation, and so on. Together with the principle of optimal design, these mutually held beliefs are the "context" for understanding her intended meaning. They allow you to make the inference that "Nice job" is intended as sarcasm rather than praise.

Rather than pursuing a technical discussion of "mutual knowledge," we discuss the more general notion of perspective taking. By *perspective taking*, we refer to people's attempts to take into account not only what they know to be mutually held beliefs, but also what they assume their partner believes from his or her perspective.

If everyone held the exact same beliefs and assumptions about the world, all perspectives would be identical and there would be no reason to assume that interlocutors should observe the principle of optimal design. Every piece of information would already be mutual by definition. The only reason to assume that interlocutors should follow this principle is because perspectives differ.

The realization that other perspectives differ from our own develops early in life. Between the ages of 3 and 4, children realize that although they might have a certain belief, others might have a different belief – which is false – about the same state of affairs. The ability of children to distinguish between their own true beliefs and others' false beliefs has been demonstrated in the now classic "false belief" paradigm (e.g., Astington, Harris, & Olson, 1988; Gopnik, & Wellman, 1992; Leslie, 1987; Perner, Leekam, & Wimmer, 1987; Wellman, 1990). In a typical false-belief task, a child is presented with a box that contains candy. The child sees the candy and also sees that Sally sees the candy in the box. Then Sally leaves. In front of the child, the experimenter then replaces the candy with pencils and closes the box so one cannot tell what it contains. Sally returns to the room. The child is asked what Sally thinks is in the box. Children younger than 3 years tend to say that Sally will think the box contains pencils. Once they develop a more adult-like understanding of the other's mind, they realize that

https://doi.org/10.1017/CBO9780511808098.010 Published online by Cambridge University Press

Sally will probably falsely believe the box has candy, not pencils. What they develop is an appreciation of differences in perspective.

Because this ability to distinguish between one's own perspective and the perspectives of others emerges early in life, it makes sense that people would use it in conversation, as the principle of optimal design suggests they should. However, the mere fact that people *have* the ability does not mean that they necessarily *use* it. In fact, in laboratory studies we find that this principle is violated across the board in the use of language – by speakers, addressees, and overhearers alike. We show that instead of following this normative principle, people use heuristic processing to resolve ambiguity.

SELF-ANCHORING AND ADJUSTMENT

When making judgments under uncertainty, people often anchor their responses to particular values – even irrelevant values – and adjust as needed (Kahneman, Slovic, & Tversky, 1982; Tversky & Kahneman, 1974; see also a review in Griffin, Gonzalez, & Varey, 2001). Such anchoring has been demonstrated in a variety of content domains with a variety of tasks. Anchoring occurs even when the exposure to the anchor value is completely incidental to the main task (e.g., Wilson, Houston, Brekke, & Etling, 1996). People show anchoring in diverse tasks, from the evaluation of utilities, to the prediction of future performance, to the attribution of causes of behavior (e.g., Gilbert & Osborne, 1989; Gilbert, Pelham, & Krull, 1988; Hershey & Schoemaker, 1985; Switzer & Sniezek, 1991; Quattrone, 1982). Anchoring coupled with insufficient adjustment typically results in a biased judgment (e.g., Slovic & Lichtenstein, 1971). Griffin and Tversky (1992), for example, used an anchoring and adjustment theory to explain why people are often over- or underconfident in their judgments. They found that people tend to anchor on the strength of evidence while insufficiently adjusting for what they know about the weight of the evidence.

Faced with the uncertainty of language, we propose that interlocutors also use such anchoring and adjustment. Instead of using information that is shared with their conversational partner, they anchor on their egocentric perspective and then attempt to adjust to their partner's perspective (Keysar, Barr, & Horton, 1998). We call this the *perspective adjustment theory* of language use. Thus, when addressees attempt to understand what speakers mean, they use an egocentric perspective. They also monitor for egocentric interpretations that violate the assumption that interpretations should rely on mutual knowledge. For example, if the egocentric interpretation uses information that is inaccessible to the speaker, the addressee detects the violation and readjusts the interpretation to align with the speaker's perspective. This adjustment, however, is often insufficient, resulting in egocentrically anchored interpretation errors. We demonstrate the value of this theory with studies on understanding and then show similar self-anchoring effects with speakers and observers.

https://doi.org/10.1017/CBO9780511808098.010 Published online by Cambridge University Press

SELF-ANCHORING WHILE UNDERSTANDING INTENTION

Consider the following situation as an illustration. Suppose you are taking a tour of the University of Chicago campus, and the guide stops next to Frank Lloyd Wright's Robie House. As the tour guide is talking about the Robie House, another building catches your attention because it is particularly unattractive, an unusual sight on that campus. At that point, you hear the guide say, "This building used to be a dormitory." In order to understand this phrase you must successfully identify the intended referent of "this building." The referent is most likely the Robie House because it is the topic of conversation and the focus of the group's attention. Indeed, when the Robie House was operated by the Chicago Theological Seminary between 1926 and 1963, it was used as a dormitory and dining hall. Assume also that you know that the guide is not looking at the other, unattractive building you just saw. One possiblity is that you interpret "This building used to be a dormitory," by considering your knowledge of what the guide is focused on, the topic, and so on. In other words, you would assume that the phrase has been optimally designed with respect to mutual knowledge. This information would immediately lead you to the conclusion that the guide is referring to the Robie House. However, the perspective adjustment theory makes a different prediction.

Instead of restricting your interpretation to what you know about the perspective of the speaker, your initial interpretation is anchored egocentrically – that is, in information available to you. This would lead you to identify quickly the unattractive building you were looking at as the intended referent of the guide's utterance, "This building used to be a dormitory." However, given that you know the perspective of the guide, you attempt to adjust to the guide's perspective. In general, then, the theory suggests a systematically egocentric error pattern, although the error might sometimes only be momentary. Such an interpretation error would occur whenever available, egocentric information suggests a different interpretation than the shared perspective.

To test between these two alternatives, Keysar, Barr, Balin, and Paek (1998) created a situation analogous to the Robie House case, except the participants in the experiment conversed about pictures of objects such as buildings, not the objects themselves. Unknown to the participants, we controlled critical moments during the experiment when we diverted the attention of the addressee to a picture that was not part of mutual focus. Immediately following this diversion, addressees heard an utterance from a confederate director that referred to the picture in mutual focus. Each critical moment was analogous to the example of looking at the unattractive building, which was not the topic of conversation, and then hearing, "This building used to be a dormitory."

We reasoned that if addressees interpret this utterance egocentrically, then the picture of the unattractive building should be identified as the referent. This temporary error should delay the identification of the picture of the Robie House as the actual referent. Indeed, participants in our experiment took longer

https://doi.org/10.1017/CBO9780511808098.010 Published online by Cambridge University Press

to identify the Robie House as the referent when their attention was diverted to the unattractive building (which also could be a referent of "this building"), compared to a control condition when they were looking at a picture of a truck (which could not be a referent of "this building.") Such interference is a direct result of the egocentric anchor that picks the unattractive building as a temporary referent of the utterance.

If understanding is guided by the principle of optimal design, then differences in perspectives should not matter. Instead, we found that differences in perspective lead to a systematic pattern of errors, even when the differences are well known. In the case of the Robie House, the unattractive building was not the topic of the conversation, but it was physically present. The perspective of the tour guide was different to the extent that the guide was not focused on that building. In one sense, perspectives always diverge as individuals have private thoughts about things that are not present. We also explored the possibility that addressees search even their private thoughts for potential referents.

Consider the use of pronouns: Pronouns can be used to make definite reference to specific entities in the world, just as the demonstrative reference in "this building." When a speaker says, "He is courageous," the speaker is using the pronoun "he" to refer to a specific individual. The use of pronouns, like other definite descriptions, presupposes that the referent is readily identifiable by the listener (Chafe, 1976). Therefore, when a speaker is using a pronoun appropriately, the referent of that pronoun should be selected from among entities that are mutually known. Because the perspective adjustment theory suggests differently, we investigated how people understand pronouns.

The experiments were modeled after the following situation: Suppose that a history professor and a student are discussing topics for a term paper. After considering several options, the student proposes to write about Niccolo Machiavelli. The mention of Machiavelli reminds the professor of John Smith, a colleague whose political intrigues had recently been revealed. As the thought of John Smith crosses the professor's mind, the student comments: "I think he has been greatly misunderstood." Who would the professor take the pronoun "he" to refer to? Our experiments show that the professor indeed interprets the utterance egocentrically, and in some cases identifies the referent of "he" to be John Smith instead of Machiavelli. The participants in our experiment made incorrect interpretations three times as often when the professor was reminded of an intrigue-related colleague whose name was John compared to a colleague whose name was Mary. In addition, when the participants correctly identified the referent of "he" to be Machiavelli, there was evidence that they had to adjust their perspective: They took longer to identify Machiavelli as the referent of "he" when they were thinking of a colleague named John rather than a colleague named Mary. So when the private thought was about a person whose gender fit the pronoun, that person was considered as the referent. The fact that the thought is private and not shared by the other person is not sufficient to preempt anchoring in an egocentric interpretation.

https://doi.org/10.1017/CBO9780511808098.010 Published online by Cambridge University Press

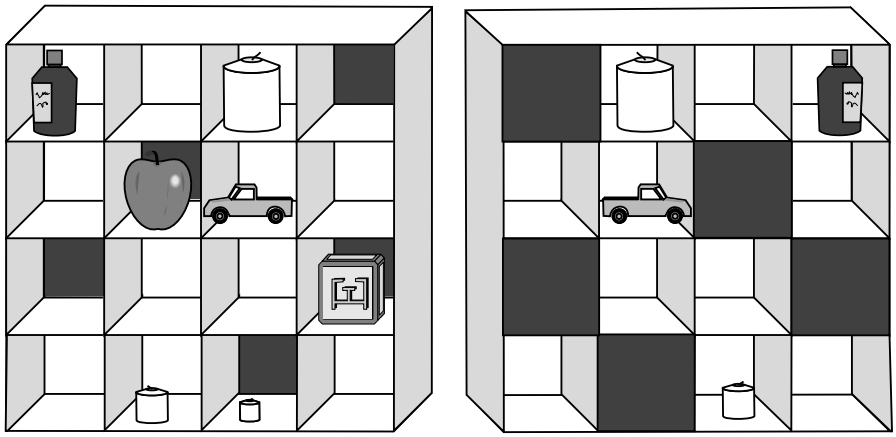

Figure 8.1. The array of objects from (A) the addressee's and (B) the director's perspectives. The two views are distinct because of the occluded slots. The critical instruction ("Put the small candle next to the truck") picks out a different candle from the director's perspective (the shared candle) than from the addressee's perspective (the occluded candle).

Perhaps the most dramatic demonstration of egocentric interpretation plus adjustment comes from an experiment that we recently conducted using an eyetracking system (Keysar, Barr, Balin, & Brauner, 2000). One participant sat in front of an array of objects as shown in Fig. 8.1. Another participant, who was actually a confederate, sat on the other side of the array. They played a simple game: The confederate (the "director") received a picture that included objects from the array but in different locations. The director then instructed the participant, the addressee, to rearrange objects to match the picture. Although most objects were mutually visible, some objects were occluded so that only the addressee could see them. They did not appear in the director's photograph and were therefore not part of the game.

Using an eyetracking device, we tracked the addressee's eye movements as they followed the director's instructions. It has been demonstrated that as soon as people identify an object as a referent, their eye fixates on that object as a precursor to the reach of the hand (Tanenhaus, Spivey-Knowlton, Eberhard, & Sedivy, 1995). Therefore, the eye movement information in our experiment indicated which objects addressees were considering as referents at any given moment.

Figure 8.1 illustrates how we tested the theory. At a critical point, the director told the addressee, "Put the small candle next to the truck." The only candles that are relevant here are those that are visible from the *director's* perspective: a large candle on the top row and a smaller one on the bottom row. Clearly,

https://doi.org/10.1017/CBO9780511808098.010 Published online by Cambridge University Press

the director is talking about the mutually visible candle on the bottom row. The addressee, however, can also see a third candle that is obviously invisible to the director. Would the addressees consider the occluded candle as the intended candle?

Our data unambiguously show that they do. Addressees take the occluded candle to be the referent before correcting themselves and adjusting to the director's perspective. In 25% of the cases, the addressee actually reached for the occluded candle, sometimes correcting in midair, sometimes lifting it briefly, returning it, then picking up the correct candle. This demonstrates strikingly how addressees egocentrically identify a referent and then adjust to the speaker's perspective.

However, even when the addressee did not reach for the occluded candle, there was evidence for an egocentric anchor and adjustment. As soon as the addressees heard the expression "the small candle," their eye was fastest to fixate on the occluded candle – suggesting that they initially identified it as the intended candle. In most cases, one could see the adjustment as the eye moved from the occluded candle and eventually fixated on the candle next to it, the one intended by the director. Such adjustment delayed the identification of the intended candle: The addressees fixated on the intended object much later when the occluded slot contained a small candle than in a control condition, when the occluded slot contained an entirely different object that could not be a referent (e.g., a small toy monkey).

The robustness of the effects, along with the fact that participants reached for objects that the director could not see at all, made us somewhat concerned. We suspected that participants might not have paid attention to the difference between visible and occluded objects. It seemed possible that participants simply did not notice which objects were actually occluded from the director's perspective. To test this possibility, we asked participants to hide the occluded objects themselves. In spite of this, the experiment demonstrated the same interference effects and the same tendency to reach for those objects that could only be referents from the participant's egocentric perspective.

These experiments demonstrate that even when people have full knowledge of the other's perspective – even when it is very clear that the director cannot see certain objects, and that the objects should not be relevant to their interpretation – addressees do not routinely use that knowledge to constrain comprehension initially. They do not assume that speakers' utterances are optimally designed to match the shared perspective. Instead, they comprehend utterances from their own perspective, and then adjust to correct any error that arises.

Self Anchoring and Adjustment in Speaking

Although addressees do not follow the principle of optimal design, speakers' behavior implies they do. People speak differently to different audiences (e.g., Fussell, & Krauss, 1989; Krauss & Fussell, 1991). For example, one study

https://doi.org/10.1017/CBO9780511808098.010 Published online by Cambridge University Press

demonstrated such "audience design" by asking New Yorkers to describe pictures of New York City to other people (Isaacs & Clark, 1987). When they were talking to other New Yorkers, they tended to use proper names such as "the Citicorp building." However, when New Yorkers addressed people who were not as familiar with the city, they were more likely to provide a longer description, such as "the building with the slanted roof." Such audience design suggests that speakers are following the principle of optimal design. Nevertheless, we discovered that speakers, like addressees, use an egocentric anchoring and adjustment mechanism.

What would seem to be an easy action – generating a sentence – requires complex mental machinery (Levelt, 1989). The production of a sentence involves several steps: planning a message (what the speaker intends to say); putting the plan into linguistic form (determining how to say it); and finally articulating it. There is also a monitoring system that triggers revisions when it detects errors.

At what point in this complex process does the speaker design the utterance for the specific addressee? One possibility is that audience design occurs at the outset, so that when speakers plan their messages, they already tailor the plan to their particular addressee. Another possibility is that planning a message is a relatively "egocentric" process – a process that does not take the other's perspective into account at all – but that the monitoring system is "in charge" of audience design. It could monitor for utterance plans that violate the perspective of the addressee, and then trigger a revision of these plans to make them more sensitive to the mutual knowledge with the addressee.

Horton and Keysar (1996) tested these two models of utterance generation by asking participants to describe various line drawings. The figures were displayed in the context of other figures, and the speakers could use the context in their descriptions. For example, when they described a circle in the context of a larger circle they could say the "small" circle. To test when audience design is employed, two kinds of contexts were used – context figures that were shared with the addressee and context figures that were privileged to the speaker. As optimal design would predict, speakers tended to make reference to the shared context more than to the context that was inaccessible to their addressee. This shows that indeed speakers in the experiment took their addressee's perspective into account when they described the figures.

Then we asked speakers to start their descriptions quickly, as soon as they saw the figures. It is important to note that the speakers were hurried only in the sense that they had to initiate their utterance immediately; they did not speak at a faster pace than normal. When speakers were hurried, their descriptions were no longer tailored to the perspective of their addressees – they used shared and privileged context to the exact same degree.

These results make sense if one assumes that before people speak, they monitor and correct their utterance plans for violations of their addressee's perspective. When speakers are under pressure, they do not have sufficient time and resources to monitor and adjust, and so they fall back on the unmonitored

https://doi.org/10.1017/CBO9780511808098.010 Published online by Cambridge University Press

utterance plans. Given that speakers did not use the perspective of their addressees when they were hurried, their descriptions reflected egocentric planning. Although this experiment shows that speakers adjust to their addressee's perspective following an egocentric plan, it does not tell us if the adjustment is sufficient. Later, we describe a study that shows that the adjustments are often insufficient.

These findings suggest an intriguing solution to the audience design goal of speaking: Speakers do not design utterances for their audience from the outset; instead, utterances seem to be tailored specifically for certain addressees only because initial utterance plans that violate the principle of optimal design are monitored, detected and adjusted. Presumably, such a process saves time and effort because in many cases there might not be a need to adjust.

THE ILLUSION OF TRANSPARENCY: SYSTEMATIC CONSTRAINTS ON COMMUNICATION

Taking the perspective of others in conversation involves interpreting an utterance from the others' perspective: Given the context available to them – what they know and believe – how would they interpret the utterance? Such assessments appear superficially easy. However, when we know the intention behind an utterance, this knowledge affects our ability to appreciate its ambiguity. Keysar (1994) has shown that when we possess knowledge of the intended meaning of a speaker's utterance, we no longer are able to accurately judge how uninformed others would perceive that same utterance. Instead, we perceive the utterance as less ambiguous than it really is – as if the intention behind it is obvious or transparent.

This phenomenon is related to hindsight bias, which has been documented in a variety of domains (Fischhoff, 1975; Fischhoff & Beyth, 1975). For example, historical events seem almost inevitable in hindsight and people claim that they could have predicted these events. However, people's ability to predict uncertain future events without the benefit of hindsight is notoriously poor. Once people know the outcome of events, they perceive the outcome as inevitable. We describe how in a similar fashion, the perception of both meaning and intention is taken as inevitable. They are taken as transparent through the utterance.

Construal and the Illusory Transparency of Meaning

Consider the case of idioms. *Idioms* are expressions whose meaning is typically not a direct function of their constituent words. Knowing the meaning of the words "kick," "the," and "bucket," and the rules of English, is not sufficient to determine that the meaning of "kick the bucket" is to die. However, not all idioms are as opaque; some feel relatively transparent (Cacciari, & Tabossi, 1993). For example, "spill the beans" seems less arbitrary because spilling corresponds to the act of revealing and the beans stand for the revealed secrets. This idiom appears to be relatively transparent because one can see a relationship between

https://doi.org/10.1017/CBO9780511808098.010 Published online by Cambridge University Press

its elements and its meaning. We suggest that this feeling of transparency is exaggerated by the very knowledge of the meaning of the idiom. Once idioms are interpreted, the meaning and the idiom's ambiguous wording become one and the same in our mind.

If an idiom seems overly transparent to us because we already know its meaning, then we should expect people who do not know the meaning to be able to figure it out. After all, it seems that the meaning is "in" the idiom. For example, native speakers of a language might expect nonnative speakers to understand their use of idioms just because their meanings seem relatively transparent.

Keysar and Bly (1995, 1999) tested the illusory transparency of meaning using a variety of archaic English idioms that are no longer familiar to English speakers. For example, the expression "the goose hangs high" used to be an English idiom that meant that the future looks good; however, most English speakers today are unacquainted with this meaning. We told some students that "the goose hangs high" meant the future looks good, and others that it meant the future looks gloomy. We then asked them to judge what uninformed overhearers would take the idiom to mean. The students who believed it meant something good predicted that uninformed peers were more likely to think the idiom meant that the future looks good rather than gloomy. The students who believed the opposite predicted that the overhearers would think that it was gloomy rather than good. Thus, the very knowledge of the intended meaning colors the ambiguous expression and makes the meaning seem more transparent than it really is.

Illusory Transparency of Intention

The illusory transparency of meaning is the feeling that the meaning of what is really a relatively opaque linguistic expression is somehow transparent. Consider now a related but different illusion of transparency. David Olson and colleagues showed that 5-year-old children are subject to the illusion of the transparency of speakers' intentions (Olson, & Torrance, 1987). In that experiment, children were told that Lucy had two pairs of red shoes: one new and one old. Then Lucy, who wanted the new pair, asked Linus to bring her "the red shoes." Linus, who did not know that she wanted the new shoes, guessed and brought the wrong pair. The children in the study were surprised that he did not bring her the shoes that she had in mind. They behaved as if the ambiguous phrase "the red shoes" uniquely identified the shoes she intended him to bring. They behaved as if Lucy's intention was transparent.

As it turns out, adults tend to behave precisely like Olson's children (Keysar, 1993; 1994). Imagine that June recommends a restaurant for dinner to her friend Mark. He goes there and hates the food. He then calls June and leaves a recorded message: "June, I just finished dinner at the restaurant you recommended, and I must say, it was marvelous, just marvelous." Mark is clearly being sarcastic,

https://doi.org/10.1017/CBO9780511808098.010 Published online by Cambridge University Press

but what would June think he meant? You know what he meant because you know he hated the dinner, but she does not have this information. You would experience an illusion of transparency if you attribute the perception of sarcasm to June because *you* know that Mark was sarcastic.

To test this, we presented people with the story about Mark and June and then played back the message he supposedly left on her answering machine (Keysar, 1994; Keysar & Baldwin, under review). However, the story was of two kinds. Mark either hated the food or he really liked it; other than that, the information in the story was identical and everyone heard exactly the same answering machine message. Although Mark had different intentions in the two versions, people's evaluation of *June's* perception should not vary. If they really take June's perspective – if they consider information that is available to *her* – then they should give the same answer in both cases. Instead, people were more likely to say that June perceived sarcasm when they knew he was being sarcastic than when they knew he was being sincere. This is precisely what one would expect if people take utterances as transparent – they behave as if the intention is obvious from what the speaker says.

As it turns out, the reason readers took Mark's intention as transparent is that he wanted June to perceive that intention. Keysar (1998) added a condition where the speaker's attitude was negative but he wanted to *conceal* it. In this case, Mark hated the restaurant, but he did not want June to know that. He wanted to spare her feelings because she recommended the restaurant. So he left her the same message, ". . . , it was marvelous, just marvelous." With this condition, participants no longer thought that June would perceive sarcasm, even though the event was negative. They were just as likely to predict that she would perceive sarcasm when he attempted to conceal his negative experience as when he had a positive experience and was truly sincere. So participants took Mark's *communicative intention* as transparent. It was as if they assumed that June would perceive whatever intention Mark wanted her to perceive.

Such an illusion of transparency could result from construal (e.g., Ross, 1990). Once we interpret an ambiguous behavior, we construe the behavior in terms of the interpretation and eventually we perceive it as unambiguous. When people knew Mark intended to be sarcastic, they reported that the repetition of "marvelous" clearly conveyed his sarcasm. When they knew he was being sincere, they reported that the repetition only underscored how much he really enjoyed it. Similarly, pauses in Mark's message were interpreted as either pregnant with sarcasm in one case or merely breathing breaks in the other. Clearly, when one construes the utterance in a particular way, it appears to include sufficient cues for the intended meaning, so that even the uninformed should be able to figure it out. The way the cognitive system construes ambiguity is a particular constraint that gives rise to an illusion of transparency in language use.

We further suggest that the illusory transparency we documented results from an anchoring and adjustment similar to the one we documented for speakers and addressees. We assume that observers anchor in the intention of

https://doi.org/10.1017/CBO9780511808098.010 Published online by Cambridge University Press

the speaker (Mark) and then insufficiently adjust to the uninformed addressee's perspective. To test this, we manipulated the amount of time participants were allotted to respond. Indeed, when participants answered at their leisure, the illusion of transparency was smaller than when they were hurried. Because adjustment takes time, the less time people have to complete the task, the more anchored they would be in the speaker's actual intention.

Similarly, the experiment provided another source of evidence for an anchoring and adjustment mechanism. It showed that responses that reflect a more complete adjustment take longer than those that are closer to the anchor points. According to an anchoring and adjustment account, if participants attribute to June the perception of sincerity when Mark is sarcastic, they must have adjusted. Indeed, addressees' response times were delayed when they indicated that June would not perceive intended sarcasm than when they indicated that she would. The opposite was true when Mark was being sincere. It took longer to attribute the perception of sarcasm to June than the perception of a sincere intent. Precisely those responses that we predict would require adjustment took longer.

Our studies converge with Gilovich, Savitsky, and Medvec's (1998) and Gilovich and Savitsky's (1999) studies that documented a related illusion of transparency – the illusion that one's own internal states, such as private preferences or emotions, are accessible to uninformed others. They showed that the illusion results from anchoring on one's own internal states and not adjusting sufficiently to the perspective of others. Unlike the illusion of transparency we demonstrated for overhearers (i.e., June), in which people thought others' intentions shone through, these studies show that people perceive their *own* internal states as relatively transparent. We now describe how speakers experience the same illusion of transparency about their own intentions.

The Speaker's Illusion of Transparency

The task of speakers is to translate their intentions into linguistic forms. To ensure they have conveyed their intentions successfully, speakers monitor their own utterances. Sometimes they recognize an utterance as ambiguous and attempt a repair or correction. However, if speakers anchor in their own intention in the process of monitoring and evaluating their own utterances, the intended meaning of their own ambiguous utterances might seem relatively transparent to them. If this is true, then the illusion of transparency is an inherent element in communication that stems from the heuristic processing of language.

One implication of the illusion of transparency is that people do not fully realize the ambiguity of their own utterances. It is particularly surprising to find such ambiguities in written text because writing can be edited. The day in September 1997 when the rock band the Rolling Stones performed in Chicago's Soldier Field, the city put up a large sign on the highway that said, "AVOID LSD TONIGHT." The city authorities were not attempting to provide advice about drug use, but instead to advise against using Lake Shore Drive in the vicinity of Soldier Field. Likewise, the owner of a pharmacy in Bethlehem, Pennsylvania,

https://doi.org/10.1017/CBO9780511808098.010 Published online by Cambridge University Press

adamantly rejected the idea that the sign above his store had dual meaning. The sign read: "Lyon's Pharmacy: We dispense with accuracy."

Part of the difficulty of writing is that when we write we are engrossed in a specific physical, emotional and mental context. It is difficult to take into account the variety of contexts that a future reader could have, contexts which might alter the meaning of what we write. More than 400 years ago, in his essay "Man Can Have No Knowledge," Montaigne described how even the same reader can interpret the same text differently on separate occasions (Frame, 1943):

> When I pick up books, I will have perceived in such-and-such a passage surpassing charms which will have struck my soul; let me come upon it another time, in vain I turn it over and over, in vain I twist it and manipulate it, to me it is a shapeless and unrecognizable mass. (p. 425)

Montaigne then goes on to describe how even the meaning of one's own text can be lost:

> Even in my own writing I do not always find again the sense of my first thought; I do not know what I meant to say, and often I get burned by correcting and putting in a new meaning, because I have lost the first one, which was better. (pp. 425–426)

Montaigne's experience demonstrates the role that his intention played at the time of writing. The text made sense to him only when he had the intention in mind, which in turn suggests that one needed the intention to make sense of the text.

Montaigne's experience might be a symptom of the way we evaluate the clarity of our own utterances. If we anchor in our own intentions when we evaluate the utterances that attempt to convey these same intentions, we might underestimate how ambiguous these utterances really are. Keysar and Henly (1998; in press) tested this idea with speakers, modeling the experiments on a study that investigated a similar phenomenon in a nonlinguistic domain (Griffin & Ross, 1991; Newton, 1990), in which participants finger-tapped a popular song so that an audience would be able to identify the song. Then they estimated how many people in the audience recognized the song. People consistently overestimated their success. Apparently, having the song in mind makes the tapping seem much less ambiguous than it really is.

In an analogous manner, might knowing our own intention when producing an utterance lead us to overestimate its effectiveness? To test this possibility, we asked people to read aloud syntactically ambiguous sentences such as, "The man is chasing a woman on a bicycle." The sentence could mean either that the man is chasing a woman who is riding a bike, or that the man is using a bike to chase the woman. We provided the speakers with a picture that disambiguated the sentence; for example, the picture showed a man running after a cycling woman. The speakers said the sentence to addressees, trying to convey the meaning expressed in the picture and then they predicted which of the two meanings the addressee would actually understand. The addressee, who did not have the picture, chose the meaning they thought the speaker intended.

https://doi.org/10.1017/CBO9780511808098.010 Published online by Cambridge University Press

In estimating the effectiveness of their utterances for addressees, there could be three possible outcomes: (1) calibrated estimation; (2) underestimation; or (3) overestimation.

Of all our speakers, only two were perfectly calibrated; one of these was an actor and the other a linguist. None of our speakers underestimated their effectiveness – they never thought that listeners did not understand them when they did. The rest of the speakers overestimated; they tended to think that they were understood when they were not. They demonstrated an illusion of transparency.

After we recorded the speakers' utterances, we asked them to judge whether they thought they would be able to understand their own intentions when they listened to the recordings the next day. The next day, when they returned to hear the recording, they were not as accurate as they had predicted. Again, they overestimated their effectiveness. They were in the same position as Montaigne when he was trying to understand his own intention.

One might want to know how often speakers experience an illusion of transparency outside the laboratory. The experiments used sentences that were syntactically ambiguous, and perhaps real-world sentences are often syntactically unambiguous. When a sentence is unambiguous, speakers cannot have an illusion of transparency by definition, because addressees can only interpret the sentence the way it was intended. In contrast to syntactic ambiguity, however, pragmatic ambiguity is the rule. *All* utterances can convey more than one intention. Therefore, the conditions in real life are even more likely to give rise to this phenomenon than the conditions in our experiments.

COMMUNICATION AND MISCOMMUNICATION OF INTENTIONS

Systematic Reasons for Miscommunication

Reasons for miscommunication are many, ranging from random error to the systematic operation of motivational and cognitive processes. Nickerson (1999) demonstrates how miscommunication can result from people's mistaken assessment of what others know and their tendency to overimpute their own knowledge to others. We have shown that even when language users are well informed about what others know, they still anchor egocentrically when taking the other's perspective. Although people might be quite good at taking into account differences in perspective when they use language, they only do so through this effortful and time-consuming process of adjustment. When people are overly busy and cognitively occupied, they might not be able to adjust sufficiently from the egocentric anchor. This has clear consequences in our overly busy world.

Inherent Limits on Adjustment

The degree of cognitive overload depends on changing circumstances; yet the studies we described indicate that even when speakers are completely at ease,

https://doi.org/10.1017/CBO9780511808098.010 Published online by Cambridge University Press

perspective taking is systematically constrained. Speakers attempting to convey their intentions anchor in those same intentions, and believe their utterances to be more transparent than they really are. Consequently, speakers might not perceive a need to make further adjustments to the addressee's perspective. In the following anecdote, the mother of the 2-year-old Ely thought that what she was saying was unambiguous:

Linda: Ely, go get me your shoes.
[Ely brings a white shoe]
Linda: Get the other shoe.
[Ely brings a black shoe]

Unlike Olson's study, where it is the child who takes Lucy's intention as transparent, in this example, it is the adult who does. For Linda, "the other shoe" had only one meaning, and her very attempt to convey this meaning inhibited her ability to perceive other possible meanings.

The Challenge of Multiple Causes

Specific instances of miscommunication are very difficult to predict. One reason is that the meaning of an utterance is often overdetermined – addressees can disambiguate utterances using multiple contextual cues. Another reason is that there are multiple root causes for miscommunication. One of the challenges is to explain how different causes of miscommunication make their unique and interactive contributions.

It is especially interesting to see how sometimes conflicting forces determine whether communication will be successful. Consider the role of motivation. When the stakes are high, a speaker is motivated to communicate effectively; however, the higher the stakes, the greater the pressure – the higher the burden on the cognitive system, leaving fewer resources for monitoring and adjustment. In the 1991 British movie "Let Him Have It," a pivotal scene tragically demonstrates how resource-depleting pressure can triumph over the motivation to communicate effectively. Derek, a basically good lad, admires and seeks to emulate his criminal friend, Chris. When the two teenagers attempt to break into a store, the police show up. An officer captures Derek and then approaches Chris, who is carrying a gun. In a tense moment, the officer demands that Chris surrender the gun. At that point Derek yells: "Let him have it, Chris!" Chris turns, shoots and wounds the officer. Later, he shoots and kills another policeman. Chris is tried for murder and found guilty. Derek is also tried for murder, because the prosecution argues that he incited Chris to shoot at the officers by yelling, "Let him have it!" The defense argues that Derek actually meant the opposite – that he meant, "Let him have the gun!" Because Chris is only 16 years old, he serves 10 years in a juvenile prison. Derek, who is 19, is convicted of murder and hanged. The movie was based on the true story of Derek Bentley.

https://doi.org/10.1017/CBO9780511808098.010 Published online by Cambridge University Press

HEURISTICS IN LANGUAGE USE

The belief that language users rely on mutual knowledge when they communicate stems from the assumption that in order to communicate successfully they *should* rely on what they know about the other's perspective. Therefore, the general assumption has been that error-free communication is the optimal and desired goal of language users, and that in order to achieve this goal, they should focus on a mutual perspective.

This normative-optimality assumption does not take into account the way the mind deals with uncertainty, of which language use is a particular case. As in the domain of decision making, the behavior of language users reflects an important tradeoff between optimality and efficiency. When solving problems, the mind takes shortcuts and uses heuristics (e.g., Tversky, & Kahneman, 1974). In this way, mental resources are conserved and solutions are reached in a short amount of time. Given that the solutions are not always optimal, they show a systematic error pattern. The benefit of such a system, though, might generally outweigh its cost – except in extreme cases such as that of Derek Bentley.

What we consistently find is that language users do not do what they "should" do according to common sense and current theory. For example, they do not routinely use readily available information about the perspective of the other. The reason is twofold. First, information about perspective is not only knowledge, it is metaknowledge: It is knowledge about who knows what. Such higher-level information typically takes more cognitive resources and more time to use. By relying on an egocentric interpretation in the case of understanding, and an egocentric plan in the case of speaking, the language user is taking a shortcut that is cheap in mental resources and relatively fast.

The second reason we rely on an egocentric process is that it is typically successful. In many cases, the overlap between the foci of the speaker and the addressee is such that an egocentric process would be sufficient for successful communication. Information that is in mutual focus is also in the egocentric focus of each interlocutor. Consider as an example the small candle in Fig. 8.1. In our experiment, we distinguished between the two perspectives by preventing the director from seeing the smallest candle. However, when people normally sit around a small table, they typically share access to all the objects that are on the table. When perspectives overlap in this way, the metaknowledge about perspective becomes superfluous. One no longer need evaluate perspective because information about who knows what would be redundant. In these cases, it is most cost effective to have a system that does not take perspective into account at all. In the typical case of perspective overlap, then, the more "expensive" monitoring and adjustment process would play no role, reaping the benefits of the quick egocentric process without paying the price of correction.

A heuristic approach to language use has another advantage. It accounts for the systematicity of errors. The normative assumption that language users

https://doi.org/10.1017/CBO9780511808098.010 Published online by Cambridge University Press

rely on mutual knowledge can only explain miscommunication as a result of random errors, noise, and unpredictable failures of attention or memory. In contrast, the perspective adjustment model predicts that when perspectives diverge or the cognitive system is overtaxed, errors will be systematically egocentric. It explains these errors as a result of insufficient adjustment. In many cases, then, miscommunication is not random and unpredictable; rather, it reflects directly the heuristic processing that underlies both how we speak and how we understand language.

BROADER IMPLICATIONS: WATCHING THE BORDER

The communication of intentions is fundamental to all social interaction. In order to navigate the social world, people are constantly inferring the motivation underlying the observable actions of others. Inferring intentions communicated through language is a ubiquitous and fundamental part of this process. Therefore, understanding the mental mechanisms of linguistic communication and understanding the conditions for their success and failure could be relevant to models in *game theory*, which is studied in a variety of the social sciences, from economics to international relations (e.g., Kreps, 1990; Laitin, 1993). A typical "game" is a situation in which the cost and benefit of people's actions depend not only on what they do, but also on what others do. For example, unsatisfied employees would want to express their concerns only if they believed that other employees were going to do the same. The risk of losing one's job would inhibit a sole complainer, but a complaint *en masse* could trigger a change in the system. In order to coordinate action, then, people can signal their intentions in a variety of ways, thereby making communication critical to people's ability to coordinate and cooperate (e.g., Chwe, 1998; Cooper, DeJong, Forsythe, & Ross, 1992, 1994).

International relations are also beset with problems of the credible communication of intentions. Communicating intentions is fundamental to the process of making decisions in international politics. In his classic book *Perception and Misperception in International Politics*, Jervis (1976) describes how two days before Germany's attack on Poland, Chamberlain sent a letter that was supposed to deter such an attack by making it clear that Britain intends to fight. The politely phrased letter was completely misunderstood as conciliatory, thus encouraging an attack.

However, such misperception of intentions occurs even when the relationship between the parties is friendly. Jervis describes how, on November 27, the Pentagon warned the commanding general in Pearl Harbor to expect "hostile action." Although the Pentagon was referring to hostile action from outside, the general interpreted it as a danger of internal sabotage. Such systematic misunderstanding can result from the way that intentions are signaled in such international games. If we understand the type of heuristics used in communication, it would allow a better understanding of how the constraints of the mind and the environment lead to successful signaling or systematic failure.

https://doi.org/10.1017/CBO9780511808098.010 Published online by Cambridge University Press