
The Effect of Common Ground on How Speakers Use Gesture and Speech to Represent Size Information

Judith Holler

Rachel Stevens

University of Manchester

Past research has investigated the impact of mutual knowledge on communication by focusing mainly on verbal communication. This study uses a wider focus, which includes speech and gesture. Speakers completed a referential communication task with recipients who did or did not share with them knowledge about the size of certain entities. The results showed that when such common ground exists between interlocutors, speakers' use of gesture and speech is affected. The main finding was that when speakers talked to recipients for whom the size information was new information, they represented this information predominantly in gesture only or in gesture and speech. However, when speakers talked to recipients with whom they shared knowledge about the entities' size, speakers encoded this information mainly verbally but not gesturally. The results are interpreted with respect to past research into common ground and language use, the pragmatics of gesture, and theories of gesture production.

Keywords: *common ground; iconic hand gestures; size information; recipient design*

Introduction

When people interact with one another in everyday talk, a multitude of social processes influence their behaviour and communication. Researchers have tried to unravel the complexities of conversational interaction using different approaches. For example, we have gained a great deal of insight from conversation analytical enquiries, particularly with regard to the orderly conduct of conversation, its sequential organisation and turn-taking systems, repairs (of self and others), and ways in which participants in conversation open their talk, end talk, invite laughter, and so on. Another approach to investigating conversational interaction is experimental research, which usually

Authors' Note: We would like to thank Howard Giles and two anonymous reviewers for their helpful comments on earlier versions of this article, Geoffrey Beattie for his valuable input through discussions of the respective topic, and the participants who took part in this study. Correspondence concerning this article should be addressed to Judith Holler, School of Psychological Sciences, Coupland Building 1, University of Manchester, Manchester M13 9PL, United Kingdom; e-mail: judith.holler@manchester.ac.uk.

involves participants interacting in the process of completing a collaborative task. One central concept that has been investigated in quite some detail is that of *common ground* (e.g., Clark, 1996; Clark & Schäfer, 1989; Clark, Schreuder, & Buttrick, 1983; Clark & Wilkes-Gibbs, 1986; Fussell & Krauss, 1989, 1992; Isaacs & Clark, 1987).

In brief, common ground can be defined as the knowledge, assumptions, and beliefs that people in interaction know (or assume) they share. When people enter a conversation, they may already share certain knowledge relating to the topic of talk, and they adjust their talk accordingly with respect to this mutual knowledge (Isaacs & Clark, 1987). This usually involves interactants being more efficient in their communication, for example, by abbreviating their speech. However, at the same time they must ensure that they communicate in a manner clear enough for their interlocutor to understand what is being said; this has been referred to as the *principle of optimal design* (Clark et al., 1983). Apart from entering conversations with some degree of common ground, people build up common ground over the course of a conversation through the contributions they make. A well-researched aspect in association to common ground is how people refer to entities and in particular how their manner of referring changes with increasing common ground. One experiment illustrating this process was carried out by Clark and Wilkes-Gibbs (1986). In this experiment, participants completed a collaborative card-sorting task (the cards showed abstract figures from a Chinese tangram game). It was found that to be able to repeatedly refer to the figures with increasing efficiency, the interactants quickly agreed on certain descriptions or labels and thus established these as common ground. For example, at the beginning of the experiment, a certain figure was referred to as “a person who’s ice skating, except they’re sticking two arms out in front”; toward the end of the experiment when common ground had been established, the same figure was referred to simply as “the ice skater” (p. 12). This example, which can be considered representative of the general change that the referential descriptions underwent with repeated use, shows that common ground can lead to more elliptical speech. Generally, utterances that are designed for addressees who share certain knowledge with the speaker tend to be less complex and rely on mutually shared understanding of meaning.

Apart from experimental investigations, examples of natural conversation support the notion that speakers in talk bear in mind what their addressees do or do not know. Goodwin (1981) introduced the terms *knowing* and *unknowing recipient* in this context to describe differences in language use and gaze patterns that speakers employ when conversing with recipients of different knowledge status, in this case recipients who are familiar with the events talked about (knowing recipients) and recipients who are not (unknowing recipients). Vygotsky (1934/1986) talked about the phenomenon of speakers anticipating recipients’ knowledge in association with his discussion of *inner speech*, or endophasy, which he claims is elliptical, like external speech can be in certain situations: “A simplified syntax, condensation, and a greatly reduced number of words characterize the tendency to predication that appears in external speech when the partners know what is going on” (p. 238; see also

Garfinkel's, 1967, analysis of "some essential features of common understandings" for further evidence for this phenomenon from natural talk).

However, these investigations into common ground and recipient designed talk in experimentally elicited and natural talk have focused on verbal communication, largely neglecting the nonverbal communicational channels that interlocutors use alongside speech. The importance of including nonlinguistic information in future analyses of common ground, and conversational interaction more generally, becomes clear if we consider a particular communicational channel, namely iconic hand gesture. Iconic gestures have been described by McNeill (1985, 1992) as imagistic representations of the hands and arms referring to concrete concepts that the speaker has in mind and intends to convey. An iconic gesture and its accompanying speech share the same semantic reference point; however, the information that is encoded in the two communicational channels is often not exactly the same, as gesture and speech tend to exhibit different semantic aspects of the idea to be expressed. An oft-cited example of an iconic gesture is the following (McNeill, 1992):¹

"she [chases him out again]"

[hand appears to swing an object through the air] (p. 13)

In this example, the speech informs us that two characters are involved in the action, that one of them is chasing the other, and that this is a recurring event. The accompanying iconic gesture, on the other hand, reveals that the chasing is done using a weapon (in this case an umbrella). Examples such as these, where the gesture carries information that is not contained in the speech, are common in spontaneous discourse and lend support to the notion that an utterance is formed by both speech *and* gesture and that therefore gesture is as much a part of language as speech (cf. Bavelas & Chovil, 2000; Clark, 1996; Kendon, 1980, 2000; McNeill, 1985, 1992).

Different theories have been proposed suggesting reasons as to why speakers make these gestures when they talk. Some researchers claim that gesturing benefits mainly the speaker himself or herself, as it helps to access lexical items in the mental lexicon during the speech-production process (e.g., Butterworth & Hadar, 1989; Hadar & Butterworth, 1997; Krauss, Chen, & Gottesman, 2000; Krauss, Morrel-Samuels, & Colasante, 1991; Rimé & Shiaratura, 1991). From this theoretical perspective, any communicative effects, if they exist, are mainly accidental. Another core theory of gesture production that considers the function of speech-accompanying imagistic hand gestures to be primarily communicative and views them as strongly motivated by aspects of the social-interactive context has been proposed. Apart from the many individual examples of gesture use observed and described by Kendon (1985, 2004) supporting this theory, there has been corroborating experimental evidence (e.g., Alibali, Heath, & Meyers, 2001; Bavelas, Kenwood, Johnson, & Phillips, 2002; Beattie & Aboudan, 1994; Holler & Beattie, 2003; Özyürek, 2000, 2002).

The intriguing question of why speakers produce gestures has spurred many researchers to investigate whether they actually communicate any meaningful

information (e.g., Graham & Argyle, 1975; Riseborough, 1981; Rogers, 1978). Perhaps the most compelling evidence to date comes from a series of fairly recent studies. Beattie and Shovelton (1999a, 1999b, 2001, 2002; see also Beattie, 2003) played video clips of individual iconic gestures produced by speakers narrating cartoon stories to independent respondents and questioned them about the information the gestures conveyed to them. Methodologically, Beattie and Shovelton used what they named a *semantic feature approach*. This involved interviewing the respondents about the information communicated by the gestures in a systematic and detailed fashion using questions that focus on a whole range of core semantic features (e.g., size, shape, movement, relative position). A robust finding emerging from these studies is that, first, iconic gestures do communicate a significant amount of semantic information both in the absence of and in addition to speech. Second, they found that particularly information about the size of entities and their relative position was communicated well by iconic gestures.

In a more recent study, Beattie and Shovelton (2006) tested whether the encoding of size information in iconic gestures is at all related to the significance of this information in the semantic context (in terms of the story line) in which it is embedded. They found that size information was encoded mainly gesturally when it was judged as highly important in the context of what was happening in the story the speakers were telling. Interestingly, and probably quite counterintuitively, the information was not represented verbally in the majority of those instances. Furthermore, when the information was considered to be of low importance, it was encoded only verbally in most of the cases. Their conclusion was that when “size matters,” it is represented predominantly in gesture only.

However, these studies leave one important question unanswered, namely whether the information contained in gestures is exclusively determined by the semantic content of a message (and the significance this information carries in the context of the story line) or whether the pragmatics of conversation also exert an influence in this respect. We therefore need to gain further insight into how speakers’ conversational goals and interactional processes influence the kind of information gestures represent and communicate.

A study by Gerwing and Bavelas (2004) has already made an important step into this direction by investigating the influence of common ground on gestural behaviour. In their experiment, one participant, who had played with a certain set of toys, had to explain to two other participants (in separate trials) what he or she did during the playing phase (without the toys being present). One of the participants who listened to the descriptions had played with the same set of toys, the other one with a different set of toys, and the speaker was aware of this. Thus, Gerwing and Bavelas created circumstances in which the recipient of the description either shared certain knowledge (i.e., about the toys) with the speaker or not. The findings showed that when speakers talk to recipients with whom they share this kind of common ground, their gestures are less complex, less precise, and less informative. Furthermore, they found that with the progression of the discourse, the information encoded in the

gestures changed or was emphasised differently with those parts of information that were newly introduced into the discourse being made more salient gesturally (e.g., by representing it more precisely). *Given* information (defined as what the listener or audience is expected to already know; Haviland & Clark, 1974, p. 512), on the other hand, tended to fade from gesture with the progression of the discourse in that the respective physical features became less salient and were often entirely eliminated from the gestures. This finding relates to literature on verbal markers of information status (e.g., Schiffrrin, 1987) as they demonstrate that gestures may be used to distinguish new from given information. Overall, the authors convincingly concluded that common ground affects not only speech but also the physical form of gestures.

These are very interesting findings, and the study provides an important stepping stone for further investigations taking a more detailed look at certain issues. In Gerwing and Bavelas's (2004) study, the influence of common ground on the physical form of gestures was examined. What we do not yet know is how common ground affects the range of semantic information encoded in gestures (i.e., how, or in which way, the gestures become "less informative" and whether any particular pattern exists that characterises this process of becoming elliptical). Moreover, we do not know how common ground influences the interplay between the two communicational channels gesture and speech. For this we need to take into consideration the verbal side of utterances. Providing first insights in these respects is the aim of the present study.

To do so, we limit our focus here to one single semantic dimension, namely size information. The reason for this is that previous studies have found that this type of information is particularly well communicated by iconic gestures, and that we already have some knowledge about how gesture and speech interact in the representation of size information (as seen above). The study described in this article uses a referential communication task that does not require speakers telling a narrative and that therefore eliminates the effects of semantic importance that information can receive from being embedded in a story line. For the successful completion of the task used here, it is necessary that speakers navigate their addressees through a picture showing a fairly busy scene (which the addressee cannot see at this point) to locate a target entity within this picture. To do so, the speaker has to refer to other entities in the picture in a clear manner, which allows the recipient to identify the entities from a large number of alternative entities reliably and without much effort. The stimulus pictures contain some particularly large entities that serve as landmarks in the process of locating the target. The crucial manipulation is that in one condition speakers are talking to addressees with whom they do not share any knowledge about the particular entities (i.e., their size). In the other condition, they talk to addressees who have seen the stimulus material beforehand and who were thus familiar with these entities and their size. Hence, making reference to the entities is equally relevant in both conditions (as the same outcome has to be achieved; i.e., correctly locating the target), but in one condition, information about the entities' size is new information, whereas in the other condition it is given as common ground does exist.

The central question this investigation tackles is whether speakers' knowledge of their addressees' familiarity with the entities' size affects how they draw on the gesture and speech modalities to communicate this information. Our assumption is that common ground does influence the semantic interplay of gesture and speech because previous research has shown that common ground affects verbal language use and other research has shown that it affects gesture. If we do find an influence of common ground on how speakers use gesture and speech to encode size information, this would have important implications for gesture production theories. However, at this stage we cannot make any precise predictions about how exactly common ground may influence gesture-speech interaction. One possibility is that when given information is omitted from speech, it is not represented gesturally because speakers may not perceive the need to communicate this information in either modality because of it being common ground. Another possibility is that the two systems somehow work around each other. For example, speakers may consider it unnecessary to provide certain information verbally, but they may represent it in gesture. This way gesture could be used as some kind of backup system (to convey information that should be considered given by the recipient but where the speaker is not certain that this is the case). Another possibility is that information that is represented gesturally in talk designed for unknowing recipients fades from gesture when common ground exists (as has been observed in previous research) but that the information is encoded verbally instead. The present investigation will provide insight into what happens verbally *and gesturally* when speakers communicate size information that either is or is not part of the interlocutors' common ground.

Method

Participants

A total of 40 students from the University of Manchester took part in the experiment in pairs. Of the 20 pairs, 4 pairs were excluded from the analysis (2 speakers were not English native speakers, 1 pair did not correctly follow the experimental instructions given to them, and, large parts of the talk of the fourth speaker were inaudible on the video recording). The remaining 16 pairs were randomly allocated to two conditions, with 8 pairs taking part in each. The pairs were mainly same-gender dyads, namely 7 out of 8 pairs (6 female dyads, 1 male dyad) in one condition (KRC) and 6 out of 8 pairs (3 male dyads, 3 female dyads) in the other condition (URC; please see Experimental Design and Procedure section for an explanation of condition labels). In both conditions, one mixed-gender dyad with a male speaker and a female recipient and one mixed-gender dyad with a female speaker and a male recipient (URC) took part.

Materials

The stimulus material consisted of five unrelated pictures from a children's book called *Where's Wally? The Wonder Book* (Handford, 1999). Each picture showed a busy scene with many different objects and characters performing various activities. Two versions were created for each of the five pictures, the original picture plus a version that was identical to the original but with the character Wally inserted into the picture. For the analysis, three pictures were chosen from the five (see Appendix A).² Recipients used sheets showing the original pictures to indicate where they believed Wally to be, based on the speaker's description (see below). The experiment was filmed split screen using two wall-mounted cameras, providing frontal views of the participants.

Experimental Design and Procedure

The study used a between-subjects design with two conditions, one which was called the knowing recipient condition (KRC) and another called the unknowing recipient condition (URC). In the URC, participants were asked to sit down on two chairs facing each other, and their choice of seats determined whether they would fulfill the role of speaker or recipient in the experiment. The positioning of the chairs provided the speaker with the view of a wall onto which the stimulus material (Version B of the pictures, showing Wally) was later projected. The recipient looked into the opposite direction and therefore could not see the projection. In the KRC, the same setup was used, but it was preceded by another phase of the experiment. This phase involved the stimulus material (Version A of the pictures, without Wally) being projected onto a different part of the wall. Both participants stood next to each other facing the projection. The speaker was asked to talk to the recipient about the picture, what it showed and what was going on, providing a good overview and summary. This way, common ground between speaker and recipient was induced in that both would have shared knowledge of the scene. After this first phase, participants were asked to sit down in their chairs again and the projected picture was swapped to the respective matching one that included Wally. From this point onward, both conditions followed the same procedure. The speaker was asked to describe to the recipient Wally's position in the picture so that they would be able to locate him correctly on a copy of the respective picture (Version A) handed to them after each description. The same procedure was repeated for each picture (the order of presentation was randomised).

Focus of Analysis

Certain entities that speakers frequently referred to when describing Wally's position and that were considered rather big in the pictorial context in that they were depicted were chosen (two knots in a pipe in Picture 1, a dome-shaped house in

Picture 2, a banana bridge in Picture 3; see Appendix A); their general perception as being big entities was validated by asking 16 independent, naïve judges to look at the stimulus material and rate the size of each of these entities on a 7-point Likert-type scale ranging from *very small* to *very big*. The average rating for the house was 5.6; for the bridge it was 4.9, for Knot A it was 5.3, and for Knot B it was 5.3, showing that they were all perceived as *fairly big* or *big*.

Coding Criteria for Speech

The descriptions of Wally's position in the pictures (which consisted of utterances such as "then go to the knot in the pipe and he's right next to it, above a multilayered cake") were coded for references made to the selected entities; each mentioning of the entities was counted as one reference (e.g., "that dome house," "a knot," "the banana bridge"). Cases in which the entities were only mentioned in verb form, such as "a knotted pipe," were excluded, as the focus of references like these is more on other entities, in this case the pipe rather than the knot (five such references occurred).

Classification of Hand Gestures

To identify and analyse gestures, McNeill's (1992) classification scheme was used. Hand gestures that represent semantic information and refer to concrete concepts are classed as iconic gestures. Rhythmical, beat-like movements following the pace and intonation of speech are called beats and do not represent any imagistic information. Pointing gestures to fictive and real objects in the surroundings of a speaker are called deictic gestures. The latter two types were excluded from the analyses, beats because they fulfill an essentially different function from imagistic hand gestures (*viz.*, the emphasis of certain parts of an utterance using rhythmical cues only) and deictic gestures because although they fulfill referential functions in talk, they do so without representing semantic information themselves. Concerning the gestures that accompanied the references made to the selected entities, only one gesture classed as a beat and one gesture classed as a deictic gesture occurred, and these two cases were excluded from the analysis. Perhaps not very surprisingly, no metaphoric gestures occurred in association with the referential task completed here. All gestures analysed in association with the references made to the selected entities were therefore iconic gestures.

Categorization of Referential Utterances

Each reference made to the selected entities was categorised as a *verbal reference* (e.g., the banana bridge), a *verbal reference + size marker* (e.g., the *big* banana bridge), a *verbal reference + gesture* (e.g., the banana bridge + hand movement tracing the bridge in the air), or a *verbal reference + size marker + gesture* (e.g., the *big*

banana bridge + a hand movement tracing the bridge). No references were made using only gesture.

Analysis of Gestures

For analysing the size information encoded in each gesture, a judgment was made using a 7-point Likert-type scale ranging from *very small* to *very big*. Each individual gesture was judged as “representing the entity as” being *very small*, *small*, *fairly small*, *neither big nor small*, *fairly big*, *big* or *very big*. One of the authors (R.S.) and an independent judge who was blind to the experimental conditions (as was the other judge) and hypotheses, but who had been shown the stimulus material and was experienced in coding gestures for the semantic information they represent, both rated each gesture on this scale. They did not hear the speech to prevent their judgments from being influenced by whether the speakers did or did not use size markers in their speech or any verbal cues indicating that their talk was designed for a knowing or an unknowing recipient.

Because exact agreement is much harder to reach using a 7-point scale than when using a considerably more limited number of scale points or coding categories, we relaxed the criterion for agreement so that directly adjacent points on the scale were considered as the judges having reached consensus (cf. Stemler, 2004). This resulted in 74% agreement. All disagreements were caused by only 2 scale points of discrepancy, except from one case where the judgments differed by 3 scale points. These disagreements were discussed, and it was found that the averages of both judges' scores for each individual gesture resolved the disagreements satisfactorily, apart from for the one case where the discrepancy in the scores was 3 scale points (for this gesture, a final score of 4 was decided because the gesture seemed to not clearly represent the entity as either big or small). The average scores were also calculated for those cases where the judges' scores were not identical but adjacent scale points.

In one part of the subsequent analysis, the judges' average size scores were used to create two categories, that is, one including gestures encoding the actual size of the entities (i.e., gestures that were scored as 5, 6, or 7) and one including gestures that did not encode the accurate size information (i.e., gestures that were scored using the scale points 1, 2, or 3). Those gestures that had received an average score of 4 (this concerned 8 gestures; 4 of these had received a score of 4 after they had been rounded accordingly), that is, less clear-cut cases, were excluded from this part of the analysis for the purpose of statistical testing.

Results

The total number of references included in the analysis was 147,³ of which 72 were used by speakers in the KRC and 75 by speakers from the URC. The following

analysis focuses on how speakers used gesture and speech to refer to the respective entities. The verbal and gestural references made were assigned to four categories (see Method section). Table 1 provides an overview of how the data distributed across these analytic categories. A G test showed that the proportional distribution of the data was significantly different in the two conditions, $G_{\text{adj}}(3, N = 147) = 15.73, p < .005$.

As can be seen from Table 1, speakers talking to knowing recipients made references mainly exclusively verbally ($55/72 = 76\%$), whereas speakers talking to unknowing recipients accompanied most of their references with gestures ($41/75 = 55\%$). Furthermore, although speakers in both conditions used size markers together with their referential expressions to an equal extent (KRC: $16/72 = 22\%$, URC: $16/75 = 21\%$), a difference emerges when this number of references is split up according to whether they were accompanied by gesture or not. References including size markers made by speakers from the KRC were predominantly purely verbal ($10/16 = 63\%$), whereas references including size markers produced by speakers in the URC were more likely to be accompanied by gesture ($12/16 = 75\%$).

Analysis of Size Information in Gesture

Apart from the frequency with which gestures were used by speakers in the two conditions, it is necessary to analyse whether these gestures actually represented accurate information about the size of the respective entities and to compare the experimental conditions in this respect. A Mann-Whitney U test was performed on the judges' size scores (see Method section) by calculating an average score for each speaker (sum of size scores of the individual gestures produced/number of gestures produced; see Appendix B for the resulting average scores). The result revealed a significant difference between the two groups ($U = 4, n_1 = 7, n_2 = 5, p < .03$), with the size scores of gestures produced in the KRC being lower than the scores in the URC. In other words, on average, the speakers talking to knowing recipients represented the entities as significantly smaller than the speakers in the URC.

For the further analysis, the gestures were categorised either as representing the respective entity accurately as big or as representing the entity "wrongly" as small (see Method section). The data are shown in Table 2. A G test showed that those gestures that did not represent the actual size of the large entities were significantly associated with the KRC, whereas those that did encode this information were associated with the URC, $G_{\text{adj}}(1, N = 50) = 17.65, p < .001$.

Of course, one could argue that perhaps an influence here is whether a respective entity has already been referred to before with a gesture. Hence, only first gestural references (see Table 3) have been taken into account to test the same association, which showed that the pattern still holds, $G_{\text{adj}}(1, N = 19) = 8.59, p < .005$.

Another possible influential factor in this respect is whether a speaker has referred to the entities in question using speech only, before making a gesture. To test

Table 1
Data Distribution Across the Four Analytic Categories for Both Conditions

	Verbal Reference			
	—	+ Size Marker	+ Gesture	+ Size Marker + Gesture
Knowing recipient condition	45	10	11	6
Unknowing recipient condition	30	4	29	12

Table 2
Frequency of Gestures That Did and Did Not Encode Size Information by Condition

	Gestures Not Encoding Size	Gestures Encoding Size
Knowing recipient condition	12	1
Unknowing recipient condition	10	27

Table 3
Frequency of First Gestural References That Did and Did Not Encode Size Information by Condition

	First Gestural References Not Encoding Size	First Gestural References Encoding Size
Knowing recipient condition	5	0
Unknowing recipient condition	4	10

this, it was not just possible to compare the two conditions in the same way taking into account only those gestures that were *actually* first references (i.e., not preceded by any verbal references to the respective entity either). This is because there would have been only one gesture in the KRC because of the existence of the “prephase,” in which participants usually referred to the respective entities. One other possibility is to look at those cases in the URC in which gestures were not preceded by any verbal (or gestural) references, which concerns 9 gestures, and compare them with those that were preceded by a verbal reference, which concerned 4 gestures. Out of the 9 gestures, 7 (78%) did encode size information, whereas 2 (22%) did not. Out of the 4 that were preceded by a verbal reference, 2 represented size information, whereas 2 did not (i.e., 50% each). This indeed very small set of data suggests that when gestures are preceded by a verbal reference to the same entity, this seems to slightly decrease the likelihood of accurate size information being represented gesturally, but not to a statistically significant extent, $G_{adj}(1, N = 13) = 0.82, ns$.

Table 4
Frequency of References Including Verbal and/or
Gestural Size Information by Condition

	Verbal Reference		
	+ Size Marker	+ Gesture Encoding Size	+ Size Marker + Gesture Encoding Size
Knowing recipient condition	10	0	1
Unknowing recipient condition	4	18	9

However, a much larger corpus of data is needed to investigate this aspect further to allow for drawing firm conclusions.

The final part of the analysis involves going back to the data distribution across the analytic categories used here; however, this time only those gestures and verbal utterances that actually did encode accurate size information are taken into account. Table 4 provides an overview of these data. It is obvious at a glance that the experimental conditions differ quite considerably.

A G test showed that this difference in the proportional distribution of the data is statistically significant, $G_{\text{adj}}(2, N = 42) = 23.85, p < .001$. If we compare individual categories, however, no significant association is found between experimental condition and whether size is encoded in gesture only (verbal reference + gesture encoding accurate size) or in gesture and speech (verbal reference + size marker + gesture encoding accurate size), $G_{\text{adj}}(1, N = 28) = 1.36, ns$.⁴ But if we compare the experimental conditions with respect to whether size information is encoded in speech only (verbal reference + size marker) or in gesture only (verbal reference + gesture encoding accurate size), a significant association is found, $G_{\text{adj}}(1, N = 32) = 21.70, p < .001$. Size is predominantly encoded verbally in the KRC and predominantly gesturally in the URC. Furthermore, to compare whether the conditions differ in terms of whether size information is encoded in speech only or in a reference that involves gesture (either as the only carrier of size information, or in addition to speech), we collapsed the categories *verbal reference + gesture encoding accurate size* and *verbal reference + size marker + gesture encoding accurate size* and found a significant association, $G_{\text{adj}}(1, N = 42) = 21.62, p < .001$. When speakers talked to unknowing recipients, they were significantly more likely to represent size in gesture or gesture and speech, whereas speakers talking to knowing recipients were more likely to represent size information verbally, without representing it gesturally.

The subsequent generic examples make clearer what this result means in terms of the actual verbal and gestural behaviour of the participants (see Figures 1-3). When conveying information about the size of the entities, speakers in the URC were much more likely to say, for example, "He's directly above the banana bridge" while moving their hand from the left side in a large upward semicircle toward the right, reaching into the

periphery of their gesture space (as in Figure 3, lower row of pictures). Or they would use the same gesture to accompany verbal utterances that also represented the size information, such as “He’s directly above the big banana bridge.” In the KRC, however, speakers would typically say “He’s directly above the big banana bridge” while moving their hand a very short way from left to right in front of the chest, drawing a very small upward semicircle in the air (as in Figure 3, upper row of pictures), therefore not conveying any accurate size information in their gesture.

Discussion

The analyses described in this article have led to a number of important findings that illuminate our understanding of how speakers in talk communicate semantic information. The crucial question was whether the encoding of size information would be affected by common ground and, if so, how this influences the semantic interaction of the two modalities, iconic gesture and speech, in the representation of this information. First, it was discovered that the pattern of how speakers use these communicational channels to refer to large entities differs significantly when they talk to knowing as compared to unknowing recipients. More precisely, using gestures to refer to the selected entities was considerably more typical for speakers talking to unknowing recipients than for speakers talking to knowing recipients. Furthermore, when speakers included size information in their verbal references, those talking to unknowing recipients tended to accompany their references with gestures; speakers talking to knowing recipients, on the other hand, used predominantly exclusively verbal references. Second, those gestures that did encode the accurate size of the large entities were mainly made by speakers who talked to unknowing recipients, and those that did not encode size information were proportionally more often made by speakers talking to knowing recipients. Overall, speakers who talked to unknowing recipients represented size information predominantly in gesture, or in gesture and speech, whereas those speakers who talked to knowing recipients represented size information predominantly in speech but hardly at all in gesture. Thus, our findings are in line with Gerwing and Bavelas’s (2004) finding that common ground influences gestural communication. At the same time, they go beyond this previous research, as we have shown how the representation of a particular semantic feature is affected and how the two modalities, gesture and speech, interact in the representation of this information.

Furthermore, the findings provide interesting insights regarding the ellipsis of speech characteristic for talk designed for knowing recipients (see Introduction). Regarding the present data, it is in fact not the case that the verbal references were condensed or contained less information when speakers talked to knowing recipients; on the contrary, the verbal references seemed to include more information (i.e., about the size of the respective entities). At first glance, this may seem slightly at

Figure 1
Gestural References Made to “The Knot” Classed as
Small (Top) and Big (Bottom)



odds with the phenomenon of elliptical speech found in previous investigations. However, taking both sides of the referential utterances (i.e., speech and gesture) into consideration leads to somewhat different insights. With respect to those cases in which the size information was represented both gesturally and verbally when speakers talked to unknowing recipients, common ground can in fact be seen as having led to more elliptical utterances, as the information was dropped from the gestures and thus represented only verbally. With respect to those cases in which size information was represented only in gesture when speakers talked to unknowing recipients, one could say that the information “swapped channels,” as it tended to be encoded verbally instead. This means that in this case the utterances do not necessarily become more elliptical (although the gestural part of them does), but the packaging of the respective information changes. This accentuates the need for considering both the linguistic and the imagistic side of utterances.

The present findings lend strong support to the theory that predicts that gestures are socially motivated in their occurrence and influenced by social processes in

Figure 2
Gestural References Made to “The House” Classified
as Small (Top) and Big (Bottom)



terms of the information they represent and the functions they fulfill. The findings thus show that those theories that explain the occurrence of gesture by linking it exclusively to speech production processes and deny any link to the speaker’s communicational intent cannot be right. Based on the present findings, we cannot (and do not intend to) exclude the possibility that gestures may also play a facilitative role in the speech production process, but we are able to conclude that this cannot be their sole function. In this sense, the findings corroborate past research (e.g., Alibali et al., 2001; Bavelas et al., 2002; Beattie & Aboudan, 1994; Holler & Beattie, 2003; Özyürek, 2000, 2002), and we can conclude with increased certainty that speakers do gesture for their interlocutors.

Furthermore, we can learn from the insights gained through this investigation that researchers interested in the effects of common ground on language use should take into account that speech and gesture *together* form an utterance. We have seen that common ground affects not only speech but also the information represented in gesture. In many cases, size information was represented in gesture but not in speech

Figure 3
Gestural References Made to “The Bridge” Classified
as Small (Top) and Big (Bottom)



and vice versa, and this semantic interaction was crucially influenced by common ground. Thus, only when considering the semantic interplay of speech and gesture can we truly discover if and how an utterance is designed for a certain recipient.

The present findings may be explained in the following way. Those speakers talking to unknowing recipients had to newly convey the size information to help create a mental image of the respective entities from scratch. Speakers talking to knowing recipients, on the other hand, were aware that the size of the entities constituted part of an existing mental representation as the recipients had previously seen the stimulus pictures. Thus, the communicational intent of the speakers talking to knowing recipients will have differed from that of the speakers in the URC. In one case, they were aiming at simply reactivating existing knowledge, whereas in the other, they aimed at helping to construct a new mental image. It appears that although the latter is preferentially done using gesture or gesture and speech, the former is preferentially done using only speech (at least with respect to the representation of size information in referential communication).

In the light of this, it seems that the words “the big bridge” (for example), when used by a speaker talking to a knowing recipient, were not primarily aimed at conveying information about the size of the entity but simply at referring effectively to something that had been seen and/or mentioned before. Haviland and Clark (1974) stated that if given information has an antecedent in memory, this allows listeners to anchor incoming information connected with this antecedent information quickly and easily. The words “the big bridge” may thus function as some kind of “verbal tag,” as it were, which would readily and reliably identify the respective entity for the knowing recipient (and allow him or her to connect new information to it, such as the location of another target in relation to the bridge). The present analysis has therefore shown that gestures may fulfill functions over and above getting semantic information across to a recipient. In line with Gerwing and Bavelas (2004), we may conclude that one other important function of gesture is to mark information status and thus to guide the recipient’s comprehension of incoming information. This kind of potential “channelling function” of certain gestures may even result in a processing advantage for the recipient; however, future research would have to confirm this assumption.

Furthermore, the results presented in this article expand on Beattie and Shovelton (2006) because they provide evidence that how size information is encoded in gesture and speech is also influenced by a speaker’s communicational intent and the particular conversational context in which talk is embedded, not only the semantic context (e.g., a story line). In the present study, conveying the size information fulfilled important communicational functions in both experimental conditions—but these appear to be quite different functions that affected the use of gesture and speech in different ways. Thus, how the speakers encoded size information was here directly influenced by the communicational context, in particular the recipients’ knowledge. To gain a more comprehensive understanding of how speakers communicate using gesture and speech, we need to take into account the semantic information a speaker tries to convey and the pragmatic functions gestures may fulfill within a particular conversation.

Of course, the study presented in this article has certain limitations. One of these is that in the present study, the participants did not verbally interact. Instead, the speakers described the scene to the extent that they believed was sufficient, which can be compared to someone taking an extended turn, for example when telling a short story or a joke. In this type of interaction, listeners usually do not interrupt and may predominantly interact with the speakers through backchannel responses and nonverbal cues. Therefore, future research will need to focus on more collaborative interaction. This would provide us with important insights into how participants in conversation work together to establish and use common ground and what role their gestural behaviour plays in this process. Another limitation of this study is that it focuses exclusively on size information. Although investigating the gestural and verbal representation of this type of semantic information was the obvious first step (for reasons stated in the Introduction), future research will have to examine other types of semantic information to provide a more comprehensive view. Currently, research that tackles both of the aforementioned points is being conducted.

Furthermore, it could be argued that the gender pairings of the participants are a potential limitation of this study because of possible differences in gestural behaviour. However, with regard to the gestural behaviour of the male dyads as compared to the female dyads, no noteworthy difference exists; considering the number of references they accompanied with gesture, the percentage values ranged from 33% to 76% for the speakers in the male dyads and from 22% to 71% for the speakers in the female dyads. On average, both male and female dyads accompanied 50% of their references made to the target entities with gestures (male dyads = 49.75%, female dyads = 49.67%). Similarly, the size information encoded gesturally varied between the scores of 1.8 to 5.6 for the male dyads (average score = 3.5) and between 2.1 and 5.7 for the female dyads (average score = 3.7). Hence, the gestural behaviour of the speakers in the female dyads is not markedly different from that of the male dyads in this experimental task. Thus, the fact that the URC had more male dyads does not appear to be a confounding factor. In the mixed-gender dyads with male speakers, the percentages of references accompanied with gestures are 17% (URC) and 23% (KRC). This may hint at a general difference in gesture use between mixed-gender and same-gender dyads, with gestures here appearing slightly less frequent in mixed-gender dyads. However, this should have had no effect on the present findings, as one male speaker talked to a female recipient in each condition (their average size scores differed considerably—7.0 and 2.0, respectively—in line with the experimental effect). Finally, in the URC a female speaker talked to a male recipient, and this gender pairing did not occur in the other condition. However, this female speaker did not gesture with the respective referential expressions, hence the part of the analysis based on gestures size scores cannot have been affected. With regard to gesture frequency, it has to be noted that three other female speakers who talked to female recipients (KRC) did not gesture with their referential expressions either. To assume that the female speaker's lack of gesturing in the mixed-gender dyad is attributable to her talking to a male recipient rather than a female recipient seems therefore not justified. However, generally, gender differences in gestural behaviour need to be more systematically addressed to allow for firmer conclusions as to whether gender does or does not have an influence on gesture use in different contexts. Previous research has yielded mixed findings, with some studies showing no significant gender differences (e.g., Duncan & Fiske, 1977; Kennedy & Camden, 1983) and others showing differences (e.g., Dovidio, Brown, Heltman, Ellyson, & Keating, 1988; Ickes & Barnes, 1977; Rosenfeld, 1966). With respect to future research, it would be particularly interesting to investigate whether gender differences in gesture use exist in the context of referential communication.

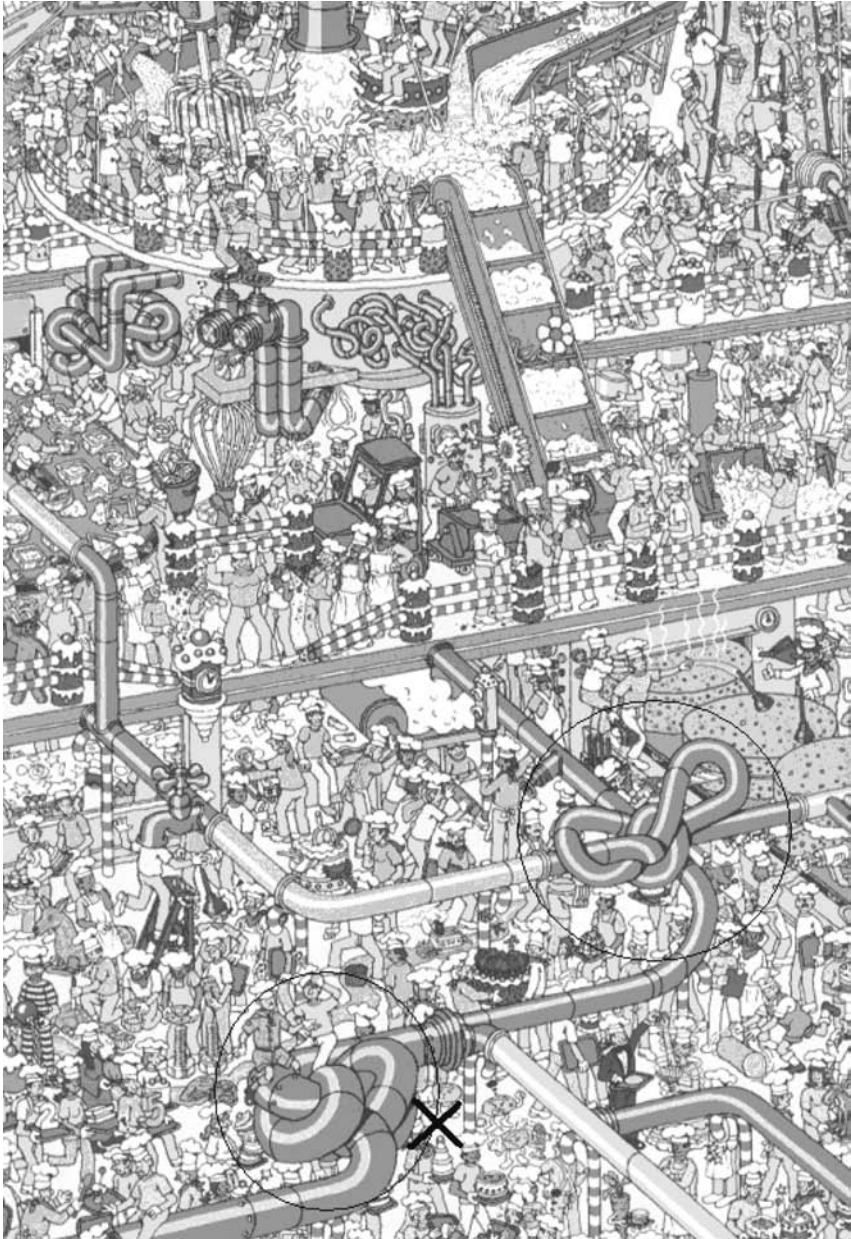
In conclusion, this article has revealed important insights into how speakers shape utterances that fit the communicational circumstances using gesture and speech. Thus, the findings have shed further light on why people gesture when they speak and how they gesture. We have also suggested how gesture may be pragmatically functioning with regard to discourse structure and discourse comprehension. These are important steps toward increasing our understanding of human communication.

Appendix A Stimulus Materials



(continued)

Appendix A (continued)



(continued)

Appendix A (continued)



Note: The black cross indicates Wally's position in the Version B pictures (see Method section). The circles refer to the target entities. Both crosses and circles are included here merely for the purpose of clarification and were not part of the original stimulus material.

Source: Illustrations © 1997 Martin Handford. From *Where's Wally? The Wonder Book*, by Martin Handford. Reproduced by permission of Walker Books Ltd, London SE11 5HJ.

Appendix B

Overview of the Average Size Scores for Each Speaker's Gestures by Condition

Participant	Average Size Score	
	Unknowing Recipient Condition	Knowing Recipient Condition
1	4.0	—
2	5.7	1.8
3	5.6	—
4	7.0	—
5	4.0	3.8
6	2.6	2.0
7	3.2	3.6
8	—	2.1

Notes

1. Speech is marked through quotation marks. The gesture is described underneath in brackets. The brackets around the spoken words mark the start and the end points of the gesture.

2. Before carrying out the experiment, we had planned to analyse material from only three pictures because of the time-consuming nature of the analysis of this kind of behavioural data and to pick those three pictures with which participants would produce the most data. Reasons for excluding the respective two pictures were that one of them showed a scene involving several cubes with letters and numbers on it; many speakers therefore simply described the target entity's position, for example, by saying, "He's above the cube with the *N* on." Regarding the second picture, the scene included very small details and entities that looked very similar so that speakers could often not make out what was depicted and were thus not able to carry out the referential task properly. Although both cases are potentially interesting with regard to gesture use (i.e., how gesture use is affected when referents are readily identifiable through verbal labels or when there is uncertainty or ambiguity about the referents), systematic manipulation and separate analyses are required to investigate them. Moreover, in the referential descriptions that the speakers did provide in association with these two pictures, they did not refer to any particularly big or small entities.

3. Henceforth, the term *references* means references to the selected entities.

4. The significance decisions for this and the following two tests are based on the Bonferroni-adjusted alpha level ($.05/3 = .017$).

References

- Alibali, M. W., Heath, D. C., & Meyers, H. J. (2001). Effects of visibility between speakers and listeners on gesture production: Some gestures are meant to be seen. *Journal of Memory and Language, 44*, 169-188.
- Bavelas, J. B., & Chovil, N. (2000). Visible acts of meaning: An integrated message model of language in face-to-face dialogue. *Journal of Language and Social Psychology, 19*, 163-194.
- Bavelas, J. B., Kenwood, C., Johnson, T., & Phillips, B. (2002). An experimental study of when and how speakers use gestures to communicate. *Gesture, 2*, 1-17.
- Beattie, G. (2003). *Visible thought: The new psychology of body language*. London: Routledge.

- Beattie, G., & Aboudan, R. (1994). Gestures, pauses and speech: An experimental investigation of the effects of changing social context on their precise temporal relationships. *Semiotica*, *99*, 239-272.
- Beattie, G., & Shovelton, H. (1999a). Do iconic hand gestures really contribute anything to the semantic information conveyed by speech? An experimental investigation. *Semiotica*, *123*, 1-30.
- Beattie, G., & Shovelton, H. (1999b). Mapping the range of information contained in the iconic hand gestures that accompany spontaneous speech. *Journal of Language and Social Psychology*, *18*, 438-462.
- Beattie, G., & Shovelton, H. (2001). An experimental investigation of the role of different types of iconic gesture in communication: A semantic feature approach. *Gesture*, *1*, 129-149.
- Beattie, G., & Shovelton, H. (2002). An experimental investigation of some properties of individual iconic gestures that mediate their communicative power. *British Journal of Psychology*, *93*, 179-192.
- Beattie, G., & Shovelton, H. (2006). When size really matters: How a single semantic feature is represented in the speech and gesture modalities. *Gesture*, *6*, 63-84.
- Butterworth, B., & Hadar, U. (1989). Gesture, speech, and computational stages: A reply to McNeill. *Psychological Review*, *96*, 168-174.
- Clark, H. H. (1996). *Using language*. Cambridge, UK: Cambridge University Press.
- Clark, H. H., & Schäfer, E. F. (1989). Contributing to discourse. *Cognitive Science*, *13*, 259-294.
- Clark, H. H., Schreuder, R., & Buttrick, S. (1983). Common ground and the understanding of demonstrative reference. *Journal of Verbal Learning and Verbal Behavior*, *22*, 245-258.
- Clark, H. H., & Wilkes-Gibbs, D. (1986). Referring as a collaborative process. *Cognition*, *22*, 1-39.
- Dovidio, J. F., Brown, C., Heltman, K., Ellyson, S. L., & Keating, C. F. (1988). Power displays between women and men in discussions of gender-linked topics: A multichannel study. *Journal of Personality and Social Psychology*, *55*, 580-587.
- Duncan, S., & Fiske, D. W. (1977). *Face-to-face interaction: Research, methods and theory*. Hillsdale, NJ: Lawrence Erlbaum.
- 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: Effects of speakers' assumptions about what others know. *Journal of Personality and Social Psychology*, *62*, 378-391.
- Garfinkel, H. (1967). *Studies in ethnomethodology*. Englewood Cliffs, NJ: Prentice Hall.
- Gerwing, J., & Bavelas, J. B. (2004). Linguistic influences on gesture's form. *Gesture*, *4*, 157-195.
- Goodwin, C. (1981). *Conversational organization: Interaction between speakers and hearers*. New York: Academic Press.
- Graham, J. A., & Argyle, M. (1975). A cross-cultural study of the communication of extra-verbal meaning by gestures. *Journal of Human Movement Studies*, *1*, 33-39.
- Hadar, U., & Butterworth, B. (1997). Iconic gestures, imagery, and word retrieval in speech. *Semiotica*, *115*, 147-172.
- Handford, M. (1999). *Where's Wally? The wonder book*. London: Walker.
- Haviland, S. E., & Clark, H. H. (1974). What's new? Acquiring new information as a process in comprehension. *Journal of Verbal Learning and Verbal Behavior*, *13*, 512-521.
- Holler, J., & Beattie, G. (2003). Pragmatic aspects of representational gestures: Do speakers use them to clarify verbal ambiguity for the listener? *Gesture*, *3*, 127-154.
- Ickes, W., & Barnes, R. (1977). The role of sex and self-monitoring in unstructured dyadic interactions. *Journal of Personality and Social Psychology*, *35*, 315-330.
- Isaacs, E. A., & Clark, H. H. (1987). References in conversations between experts and novices. *Journal of Experimental Psychology: General*, *116*, 26-37.
- Kendon, A. (1980). Gesticulation and speech: Two aspects of the process of utterance. In M. Ritchie Key (Ed.), *The relationship of verbal and nonverbal communication* (pp. 207-227). The Hague, Netherlands: Mouton.

- Kendon, A. (1985). Some uses of gesture. In D. Tannen & M. Saville-Troike (Eds.), *Perspectives on silence* (pp. 215-234). Norwood, NJ: Ablex.
- Kendon, A. (2000). Language and gesture: Unity or duality? In D. McNeill (Ed.), *Language and gesture* (pp. 47-63). Cambridge, UK: Cambridge University Press.
- Kendon, A. (2004). *Gesture: Visible action as utterance*. Cambridge, UK: Cambridge University Press.
- Kennedy, C. W., & Camden, C. (1983). Interruptions and nonverbal gender differences. *Journal of Nonverbal Behavior*, 8, 91-108.
- Krauss, R. M., Chen, Y., & Gottesman, R. F. (2000). Lexical gestures and lexical retrieval: A process model. In D. McNeill (Ed.), *Language and gesture* (pp. 261-283). Cambridge, UK: Cambridge University Press.
- Krauss, R. M., Morrel-Samuels, P., & Colasante, C. (1991). Do conversational hand gestures communicate? *Journal of Personality and Social Psychology*, 61, 743-754.
- McNeill, D. (1985). So you think gestures are nonverbal? *Psychological Review*, 92, 350-371.
- McNeill, D. (1992). *Hand and mind: What gestures reveal about thought*. Chicago: University of Chicago Press.
- Özyürek, A. (2000). The influence of addressee location on spatial language and representational gestures of direction. In D. McNeill (Ed.), *Language and gesture* (pp. 64-83). Cambridge, UK: Cambridge University Press.
- Özyürek, A. (2002). Do speakers design their co-speech gestures for their addressees? The effects of addressee location on representational gestures. *Journal of Memory and Language*, 46, 688-704.
- Rimé, B., & Schiaratura, L. (1991). Gesture and speech. In R. S. Feldman & B. Rimé (Eds.), *Fundamentals of nonverbal behaviour* (pp. 239-281). Cambridge, UK: Cambridge University Press.
- Riseborough, M. G. (1981). Physiographic gestures as decoding facilitators: Three experiments exploring a neglected facet of communication. *Journal of Nonverbal Behavior*, 5, 172-183.
- Rogers, W. T. (1978). The contribution of kinesic illustrators toward the comprehension of verbal behaviors within utterances. *Human Communication Research*, 5, 54-62.
- Rosenfeld, H. M. (1966). Approval-seeking and approval-inducing functions of verbal and nonverbal responses in the dyad. *Journal of Personality and Social Psychology*, 4, 597-605.
- Schiffrin, D. (1987). *Discourse markers*. Cambridge, UK: Cambridge University Press.
- Stemler, S. (2004). A comparison of consensus, consistency, and measurement approaches to estimating interrater reliability. *Practical Assessment, Research and Evaluation*, 9, 4. Retrieved June 1, 2006, from <http://PAREonline.net/getvn.asp?v=9&n=4>
- Vygotsky, L. S. (1986). *Thought and language* (A. Kozulin, Trans. & Ed.). London: MIT Press. (Original work published 1934)

Judith Holler is a lecturer in psychology at the University of Manchester. Her main research interest is the uses of speech-accompanying gestures in talk, the semantic interaction of gesture and speech, and the communicational effects of gestures.

Rachel Stevens is a psychology graduate from the University of Manchester. She is interested in the effect of common ground on iconic gestures.