

Speakers' use of 'action' and 'entity' gestures with definite and indefinite references

Running title: Common ground and gesture

Katie Wilkin^a & Judith Holler^{a, b}

^aUniversity of Manchester, U.K.

^bMax Planck Institute for Psycholinguistics, The Netherlands

Abstract

Common ground is an essential prerequisite for coordination in social interaction, including language use. When referring back to a referent in discourse, this referent is 'given information' and therefore in the interactants' common ground. When a referent is being referred to for the first time, a speaker introduces 'new information'. The analyses reported here are on gestures that accompany such references when they include definite and indefinite grammatical determiners. The main finding from these analyses is that referents referred to by definite and indefinite articles were equally often accompanied by gesture, but speakers tended to accompany definite references with gestures focusing on action information and indefinite references with gestures focusing on entity information. The

findings suggest that speakers use speech and gesture *together* to design utterances appropriate for speakers with whom they share common ground.

Key words: common ground, new and given information, definite and indefinite references, iconic gestures, deictic gestures, entity information, action information, ellipsis

Introduction

One of the central questions gesture researchers have tried to answer in recent years is why we gesture when we speak. This research has led to a greater understanding of the functions of co-speech gestures, and the empirical evidence suggests that they may indeed fulfil a range of quite different functions. For example, co-speech gestures appear to aid the speaker's cognition, such as the processes involved in lexical retrieval (e.g., Pine, Bird & Kirk, 2007) or conceptual planning (e.g., Hostetter, Alibali & Kita, 2007). Others argue that gestures fulfil communicative functions (e.g., Bavelas & Chovil, 2000; Kendon, 2004). For example, we know that social context in the form of visibility between speaker and addressee influences gesture rate (Alibali, Heath & Myers, 2001; Bavelas, Kenwood, Johnson & Phillips, 2002) as well as aspects of gesture form (Bavelas, Gerwing, Sutton & Prevost, 2008; Gullberg, 2006), and that addressee location can influence speakers' use of gesture space to represent semantic information (Furuyama, 2000; Özyürek, 2002).

Apart from these overt, physical aspects of the social situation, there is also evidence that more covert processes influence gestural communication, such as the interactants' thinking and understanding. Holler and Beattie (2003b) found that speakers use co-speech gestures to clarify lexical ambiguities for their addressees, both in dialogue-like interactions as well as in more monologue-like narratives. Because verbal ambiguity can be a problem for the addressee but is rarely a problem for the speaker him- or herself, these studies provide evidence that speakers do gesture for their recipient and that they take their addressees' thinking into account when gesturing. Recent research has shown that this conclusion is not restricted to the context of lexical ambiguity but that it generalises to other domains. Some of this research has focused on an aspect fundamental to successful communication, namely the knowledge, beliefs and assumptions interactants mutually share, which has been referred to as 'common ground' (e.g., Clark 1996). Studies examining verbal communication have revealed that common ground leads to more elliptical speech (e.g., Clark & Wilkes-Gibbs, 1986; Fussell & Krauss, 1989; Isaacs & Clark, 1987), amongst other things. Recently, researchers have started to investigate the effects of common ground on gesture use. Gerwing and Bavelas (2004, Study 1) showed that speakers used less complex, precise and informative gestures when they talked to addressees with whom they shared common ground than when talking to addressees with whom they did not share common ground. Similarly, Holler and Stevens (2007) found that speakers encoded less information about the size of entities in gesture when their addressees shared common ground with them regarding this semantic aspect than when they did not. Similarly, Parrill (2010) found that speakers encoded significantly less information about the ground element of an event they were describing when they mutually

shared knowledge about this event with their interlocutor than when they did not. Further, findings by Jacobs and Garnham's (2007) suggest that speakers gesture at a lower rate when common ground is built up based on repeated narrations of the same story to the same listener (see also Holler, 2003). Taken together, this evidence may lead us to conclude that gestures, like speech, are more elliptical when common ground exists. On the other hand, a study by Holler and Wilkin (2009) revealed that speakers in their common ground condition gestured at a *higher* rate when common ground existed and that they encoded statistically as much semantic information in their gestures in this condition as in the one without common ground.

Several factors could explain the discrepancies between these findings. For example, the studies differed in the way the participants were interacting during the task (free vs. restricted interaction), and in the type of tasks the participants completed (e.g., narratives vs. referential communication tasks). Studies systematically investigating these and other potential factors are currently underway. What we can conclude to date is that common ground appears to influence gestures in a variety of different ways and that the semantic interplay between gesture and speech in this context does not seem to be characterised by one simple pattern. Further research is needed to arrive at a more complete view of how common ground influences communication.

The present study focuses on utterances including definite and indefinite references, and amongst those on references including an indefinite article ('a' or 'an') or a definite one ('the'). Such articles mark information either as 'new' or 'given'. There has been some variation in terms of how new and given information has been defined; in the light of this, Prince (1981) has established three different notions of 'givenness'. This includes the

notion of ‘givenness’ as predictability of a lexical item in its sentential context (based on, for example, Halliday, 1967 and Kuno, 1972), ‘givenness’ as saliency in terms of an entity being in the addressee’s consciousness (based on Chafe, 1976), and ‘givenness’ as shared knowledge – knowledge the speaker assumes their addressee knows, believes or is able to infer (based on Clark & Haviland, 1977). In the present article, we use the latter definition. Consequently, ‘new information’ is here defined as that which the speaker believes is not yet known by the addressee (i.e., information which is not yet part of the interlocutors’ common ground).

Past research has focused on how given information is communicated in discourse and how speakers lexically mark such common ground (e.g., Fetzer & Fischer, 2007); however, little research in this area has focused on gesture. One exception is a study by Gerwing and Bavelas (2004, Study 2). This study included an analysis of ten dialogues in which one person had played with a particular toy and described this toy and the actions carried out with it to another person who had not played with or seen the toy. Thus, initial references to features of the toy and its actions were new information, with subsequent information of this kind being given information. Their gestural analysis showed that the accumulating common ground did influence the form of the gestures in that given information was made less salient gesturally and gestures accompanying given information were smaller and less precise. Levy and McNeill (1992) as well as McNeill, Cassell and McCullough (1993) have analysed speakers’ verbal and gestural repeated references to the same characters in a story. Their focus was on pointing gestures accompanying initial and subsequent, more attenuated references (mainly in the form of pronouns and zero

anaphoras). Pointing gestures were found to occur less frequently with attenuated references (i.e., when the information was given).

The present study compares speakers' gesture use with definite references (e.g., including the lexical marker 'the') and indefinite references (e.g., including the lexical marker 'a/an') in terms of gesture rate and the type of gestures used. The analyses aim to further explore how speakers communicate given and new information in speech and co-speech gesture, going beyond previous research by focusing on grammatical articles (rather than pronouns) and on iconic as well as deictic gestures. Due to the incoherent picture emerging from the previous studies into common ground and gesture, no firm predictions regarding the pattern we may observe can be made. The data used in the analyses stem from an experiment which was originally designed to manipulate the amount of common ground that exists from the outset of a conversation (common ground based on prior physical co-presence, Clark & Marshall, 1981). Participants took part in pairs, with one speaker being allocated the speaker role (and the other the role of the addressee); this person later narrated a story they had seen on video to the addressee participant. In the 'no common ground condition' (NCG), the addressee participant had no knowledge about the story prior to the speaker's narrative. In the common ground condition (CG), the addressee participant watched individual scenes from the video together with the other participant (who then watched the entire video, on their own, prior to narrating the full story to their addressee). For the present analysis we collapse the data from both the 'common ground' and the 'no common ground condition' as they are equally suited to examine common ground that accumulates during the course of a narrative (common ground based on

linguistic co-presence, Clark & Marshall 1981). However, we also use the original experimental common ground manipulation as a variable in some of the analyses.

Method

Experimental Design

The present study was conducted as an additional analysis on a subset of the data published in Holler and Wilkin (2009).¹ It is based on a between-subjects design with two conditions: the ‘common ground’ condition (CG), in which participants shared some experimentally induced knowledge about the stimulus material, and the ‘no common ground’ condition (NCG), in which participants did not share any experimentally induced common ground (other than that which accumulated during the narrative).

Participants

The present analyses are based on fifty-six students (22 female and 34 male) from the University of Manchester who took part in the experiment (all received either payment or experimental credits for their participation). All individuals were right handed (as measured by the Edinburgh Handedness Inventory, Oldfield, 1971) and native English speakers. Each

¹ Only a subset was used in the present analysis because the data in Holler and Wilkin’s (2009) study were analysed in two steps, the first one focusing on a smaller subset, at which point the present analysis was conducted.

participant was allocated to a same-sex pairing, which was then randomly assigned to one of the two experimental conditions resulting in 14 same-sex pairs in each condition.

Materials

A short (about 7 minute long) video was used as the stimulus material. It contained a story in which child and adult human characters were involved in different everyday activities, such as mending a car, grocery shopping, or playing in a barn. From this video, six short scenes (each 2-5 seconds in length) were selected for the common ground manipulation (see Procedure). The participants were filmed in a social observation laboratory including two high definition wall-mounted cameras, each providing the view of one participant, feeding into a dvd recorder in a split-screen format.

Procedure

In both the CG and the NCG conditions, two participants took part at a time, allocated to the roles of speaker and addressee based on their seat choice. The speaker watched the six selected scenes, followed by the whole video. However, in the CG condition, the addressee watched the six scenes together with the speaker (but was absent while the speaker watched the full video). During the following narration phase, the participants sat opposite each other, and the speaker was instructed to tell their addressee what happened in the story as a whole, bearing in mind that a) their addressee did not know anything about the story (NCG condition), or b) that their addressee already shared some knowledge about the story with

them (CG condition). Addressees were told before the experiment that they would be asked content-related questions at the end. They were also told that they were free to signal their understanding during the narration as they felt appropriate, but that they should not interrupt the speaker to ask questions.

Analysis

Participants' gestural and verbal behaviour relating to five of the six selected scenes was included in the analyses. The sixth scene was excluded due to similarity with another part of the video, which made it impossible to decide for certain in all instances which of the two events in the story speakers were referring to.

Speech segmentation

All descriptions of the five target scenes were transcribed verbatim. To identify the respective parts in the narratives, each event was defined in terms of what it comprised semantically (i.e., ideational units, see Butterworth, 1975; Holler & Beattie, 2002). Only those parts of the narratives were analysed that included semantic information from the five target scenes. The percentage agreement between two independent coders identifying the first and the last word to be considered part of the scene was 87.6%. All discrepancies were resolved through discussion.

Coding for definite and indefinite references based on grammatical determiner

Within the individual speech segments, the following determiners were identified (including both grammatical articles and demonstratives): ‘the’, ‘that’, ‘a’, ‘an’, and ‘this’. ‘The’ and ‘that’ were both regarded as lexical markers of common ground (or given information) and were therefore combined in the analyses. We are here not referring to ‘that’ being used as a demonstrative singling out a referent in physical space (such as when pointing to something) but as a demonstrative in the absence of any nonverbal or physical contextualising cues; an example would be the utterance ‘and then that light blue car came along’ to refer back to a scene the interlocutors had seen together, or to the car when they had mentioned it beforehand. Similarly, ‘a/an’, and ‘this’ were combined as markers of no common ground (or new information); again, we are here referring to the demonstrative ‘this’ being used without any contextualising information (such as a pointing gesture to an object in the physical surroundings), but, rather, as a general determiner for a referent outside of the common ground, as in ‘suddenly, this car comes around the corner’ to refer to a car which is not present at that moment. That is, ‘the’ and ‘that’ are here classed as definite references and ‘a’/‘an’ and ‘this’ as indefinite references. While the terms definite and indefinite references also refer to anaphoric expressions (Keysar, Barr, Balin & Paek, 1998), the present analysis limits its focus to references including basic definite and indefinite determiners. This means that when a gesture accompanied a part of speech that contained both a grammatical article and an anaphora we used the grammatical article for classification (see Examples 1 and 2). The rationale for this decision was that, while previous analyses have focused on gestures accompanying attenuated references to characters in the form of zero anaphoras or pronouns (e.g., Levy & McNeill 1992; McNeill, Cassell & Levy, 1993), in our data the use of pronouns was not that prevalent; instead, most

references to the scenes constituting the analytic focus included the entities' grammatical articles (+ noun). The present analysis therefore complements those earlier studies.

Gesture coding

Gesture category. Co-speech gestures were identified and categorised according to McNeill's (1992) categorisation scheme by coding them as iconic, metaphoric, deictic (in the present data only abstract deictics occurred), or beats, complemented by Bavelas, Chovil, Lawrie and Wade's (1992) category of interactive gestures². The percentage agreement between two independent judges using these categories classifying all gestures co-occurring with references to the five target scenes was 79.9%. Again, all discrepancies were discussed and resolved.

Gesture type. For the second part of the analysis, all iconic and deictic gestures were further classified as 'action gestures' (e.g., an iconic gesture representing someone picking something up; an iconic gesture performed with a single finger moving from left to right to indicate a car driving past), or as 'entity' gestures (e.g., a deictic gesture indicating the presence of an entity, or an iconic gesture representing a whole or part of an object, such as by using the index fingers to outline the square shape of a window). These examples illustrate that the distinction between entity and action gestures is not an absolute one —

² The categories of 'beats' and 'interactive gestures' seem to overlap (cf. Bavelas, Chovil, Coates & Roe, 1995; Jacobs & Garnham, 2007), but during our coding procedure we encountered some gestures that we felt clearly belonged to one and not the other class of gestures, based on the form criteria described by McNeill (1992) and Bavelas et al. (1992); we therefore included two separate categories to capture these gestures.

gestures classed as action gestures included those that were considered to be *primarily* encoding information about an action, but may have included information about entities (such as the narrator's hand carrying out the action representing the character's hand); the rationale for calling these 'action gestures' was that they seemed to *foreground* the action component of the gestural representation. Gestures classed as 'entity gestures' always encoded just entity information. The inter-observer reliability of two independent coders for this binary categorisation was 94.3%. The few disagreements that occurred were subsequently resolved through discussion.

Examples (1) and (2), and the following description, illustrate the coding of one speech segment and its accompanying iconic and deictic hand gestures. The underlined words are the definite and indefinite references based on determiner, and the square brackets mark individual gestures, indicated as subscript preceding the respective gesture and numbered consecutively. The superscript letters within each square bracket indicate whether the gesture primarily encoded action information (A) or entity information (E). If an article type was not accompanied by a gesture, it was coded as having no accompanying gesture (subscript N).

(1)

G₁[the boy^E] G₂[picks up the piece^A] of litter, G₃[and puts it in the bin^A]

G1: abstract deictic gesture pointing towards the right hand side of the gesture space, referring to the boy.

G2: iconic gesture showing someone grabbing something which is moved upwards (palm pointing downwards).

G3: iconic gesture showing someone holding something enclosed in the hand which moves down and forwards, stopping at about chest height in front of the speaker's body.

(2)

NThe kid... G1[picks up a bit of litter^A] off Nthe floor G2[and puts it in a^A], G3[in a litter bin^E] G4[which is a little basket^E] G5[attached to a lamppost^E]

G1: iconic gesture showing someone grabbing something which is moved upwards (palm pointing downwards).

G2: iconic gesture showing someone holding something in the hand which moves down and sideward, stopping at about lap / thigh height to the side of the speaker's body.

G3: iconic gesture showing the vertical, straight sides of a small, imaginary, upright container.

G4: iconic gesture showing the sides and the base of a small, imaginary, upright container.

G5: iconic gesture showing the narrow width, elongated shape, and vertical orientation of an imaginary object

If more than one gesture accompanied a stretch of speech that contained only one article type, then the gesture performed closest in time to the respective determiner (i.e., the gesture with the strongest temporal relation to the word, 'the', 'that', 'a/an', or 'this') was counted. Furthermore, if a part of speech containing an article type had no gestural accompaniment, while a subsequent gesture performed in synchrony with an immediately following part of speech nevertheless appeared semantically related to the preceding speech segment, this gesture was not counted as an accompanying gesture for the former article type but for the one it co-occurred with. Thus, temporal co-occurrence rather than semantic relation was used as the main criterion (although this was equivalent in most cases).

Results

The analyses reported here are based on a corpus of 277 references including the respective grammatical articles. For the statistical analyses, an alpha level of .05 is used throughout (all tests reported are two-tailed).

Definite and indefinite references

Across both conditions, references including definite determiners, ‘the’ and ‘that’, were used more frequently (180 times in total) when compared with references including indefinite determiners, ‘a/an’ and ‘this’ (97 times in total). This is not surprising since we took into consideration the first time an entity was being referred to, as well as all subsequent references, and speakers tended to refer to some of the entities repeatedly (such as the characters involved in the storyline) – thus establishing exactly the sort of common ground we intended to capture.

Co-speech gestures

Based on our corpus of 277 references and 210 co-speech gestures, we then focused on the proportion of gestures accompanying each reference type (i.e., number of gestures/number of definite references or indefinite references), Table 1. Firstly, the analysis revealed that the same proportional amount of references classed as indefinite was accompanied by gestures as references classed as definite ($z = 0.329$, N -ties = 24, $p = .742$, ns). This pattern held when we considered the individual gesture categories separately, with the exception of iconic gestures, of which a higher proportion accompanied definite references ($Median = 1$, $Range = 1$) than indefinite references ($Median = .75$, $Range = 1$), $z = 2.32$, N -ties = 24, $p = .021$.

INSERT TABLE 1 HERE

'Action' and 'entity' co-speech gestures

Iconic and deictic gestures that accompanied the definite and indefinite references (192 gestures in total) were classified as either 'entity' or 'action' gestures (see Method). The frequencies and percentages can be found in Table 2.

INSERT TABLE 2 HERE

A 2 (gesture type: action vs. entity) x 2 (reference type: definite vs. indefinite) repeated measures ANOVA was carried out and revealed that there was a main effect of reference type ($F(1, 27) = 4.50, p = .043$); out of those references that were accompanied by gesture, more were definite ones than indefinite ones. The main effect of gesture type was not significant ($F(1, 27) = 3.16, p = .087, ns$), meaning that, overall, speakers used as many gestures that focused on actions as gestures that focused on entities. However, the interaction between gesture type and reference type was significant ($F(1, 27) = 5.36, p =$

.028), with more ‘entity gestures’ accompanying indefinite references, and more ‘action gestures’ accompanying definite references.

When considering the experimental common ground manipulation as a third factor with two levels, CG and NCG (see Figure 1), in addition to the effects mentioned above, the statistical analysis revealed a significant interaction between common ground and gesture type ($F(1, 26) = 5.16, p = .032$), with speakers in the CG condition using mainly ‘action’ gestures, and speakers in the NCG condition using mainly ‘entity’ gestures.

However, the interaction between the common ground manipulation and reference type was not significant ($F(1, 26) = 1.13, p = .297, ns$), and neither was the three-way interaction between common ground, reference type and gesture type ($F(1, 26) = 2.37, p = .136, ns$).

Table 3 shows the association between gesture type and experimental condition when considering just those two variables.

INSERT FIGURE 1 HERE

INSERT TABLE 3 HERE

Discussion

The present analyses yielded a number of important findings. Firstly, but not surprisingly, we found that speakers used more definite references than indefinite references because they tended to refer to the same characters or objects more than once (and for half of the participants the referents were already in their common ground due to the experimental manipulation). Secondly, the findings show that speakers accompanied these two different types of references with gesture statistically equally often. However, a further analysis revealed that when splitting the amount of gestures up according to different gesture categories, speakers accompanied a higher proportion of definite references with iconic gestures than they did indefinite references. Another analysis distinguished between what we called different ‘gesture types’, which referred to gestures foregrounding information about entities (‘entity gestures’) and gestures foregrounding information about actions (‘action gestures’). This analysis revealed that ‘action gestures’ accompanied mainly definite references, and ‘entity gestures’ mainly indefinite ones. Finally, we found that the manipulation of common ground that exists from the outset of a conversation (that is, common ground based on prior physical co-presence, Clark & Marshall, 1981) interacted with gesture type; whereas speakers in the common ground condition used mainly ‘action gestures’ when referring to those segments of the story constituting common ground, speakers in the no common ground condition used mainly ‘entity gestures’ with references

to the same semantic events. Taken together, these findings suggest that common ground was associated mainly with iconic gestures and action information, and no common ground mainly with abstract deictic gestures and entity information. The main conclusion to be drawn from these findings is that the semantic interplay between gesture and speech is not characterised by a simple, parallel pattern according to which both speech and gesture are more elliptical in the context of common ground. Rather, it appears that speakers employ the two modalities to package the information they intend to convey in a manner most appropriate with respect to the recipient's knowledge status, which can involve more complex representations in gesture even when common ground exists.

This appears to fit the results obtained from an earlier analysis of a similar dataset (Holler & Wilkin, 2009). Amongst others, this analysis revealed that speakers gestured at a higher rate (with regard to iconic and deictic gestures) when common ground existed (referring to common ground existing from the outset). Further, their findings showed that, overall, speakers' gestures did not decrease significantly in semantic content when common ground did exist as compared to when it did not. The authors argued that this does not mean that the gestures were not 'recipient designed' (Sacks, Schegloff & Jefferson, 1974). Rather, they suggest that the gestures continued to play an important communicational role, but that this role may be different to that of the gestures accompanying the same event descriptions when no common ground existed. The pattern revealed by the present analysis fits this notion; it suggests a shift in semantic focus regarding the gestural representations accompanying references to entities of different information status. The pattern is characterised by more semantically complex gestures accompanying references to information that is in common ground. Although we did not systematically quantify the

information contained in the gestures using our ‘entity’/ ‘action’ distinction, we observed that many ‘action gestures’ encoded also some entity information, whereas the ‘entity gestures’ only ever encoded entity information. Of course, the entity gestures could have been encoding information about several entities at once, and action gestures might have been highlighting just one particular dimension of a movement (e.g., direction). Due to this we cannot claim that action gestures *always* contained more information than entity gestures, but a large number of them appeared to do so. The shift in gestural focus observed in the present dataset may be one factor that could explain the lack of a difference in the amount of semantic information represented in gesture found by Holler and Wilkin (2009). The authors speculated that many of the gestures referring to information in common ground were semantically complex instead of elliptical so that they could fulfil a back-up function in case of speakers’ uncertainty about specific information being in common ground or not (i.e., with the gestural information compensating for ellipsis in speech in case it is needed). Another possibility they mention is that these fairly complex gestures may assist speakers in focussing their addressees on the correct aspect of their mutually shared knowledge. The fact that, in the present study, when entity information was in the common ground, speakers put less emphasis on the individual characters and more emphasis on the actions carried out by these is agreeable with both of these possible explanations. More research is needed to illuminate this issue further.

Our findings are in line with Foraker and Goldin-Meadow (2007) who found that speakers tend to use gesture to depict the identity of a referent when this referent is newly introduced in speech, but that they used gesture to represent supplementary information about the referent when the referent had already been mentioned.

Our findings may be conceived of as complementing those studies providing evidence of increased ellipsis in gesture in the context of given information. Levy and McNeill (1992) and McNeill, Cassell and Levy (1993) found that pointing gestures (abstract deixis) occurred mainly with initial references to characters in a narration and less frequently with later ones. Once information about the identity of the referent was in the interactants' common ground, they used no gestures with their referring expressions, or the pointing gestures were replaced by other types of gestures (McNeill et al., 1993, p.16). Similarly, Gerwing and Bavelas (2004, Study 2) found evidence of a reduction in semantic content in gesture when these gestures were referring to given instead of new information. Although our data show no statistically significant reduction in gesture use when common ground existed, they do show that speakers gesturally emphasise different semantic event aspects (i.e., the actions rather than the entities) and that they used mainly iconic rather than deictic gestures to do so. With regard to depicting entity information, the gestures in our corpus did tend to become more elliptical. However, our data throw a different light on the topic, as we have provided evidence that gestures do not always become more elliptical overall when information is in common ground (although this *can* happen), or that they simply disappear altogether because communication may be conceived of as easier when information is given. Instead, our data suggest that gestures continue to be important for communication and that they combine with speech in a variety of ways to achieve a successful and pragmatically appropriate exchange of information.

In addition to exploring the exact functions co-speech gestures fulfil in this context, future studies will need to establish to what extent the interaction patterns the present and previous studies have revealed are specific to the particular communicative situation

examined; that is, the functions of gestures may be specific with respect to whether speakers communicate information that is common ground based on prior physical co-presence, linguistic co-presence, or visual co-presence, for example. Further insights may also be gleaned from more detailed analyses which take into account the structure of individual utterances and the sequence of the gestures accompanying them; after all, the present analyses are based on aggregate data, summarising the occurrences of different gesture types across the discourse, which provides us with merely a first glimpse of what may be going on. Nevertheless, one important conclusion we can draw from the present findings is that the way in which interlocutors communicate information that is in the common ground they share appears to be complex, with partly parallel, partly complementary changes happening to gesture and speech. Thus, only a multi-modal enquiry will be able to provide us with a more complete view of communication in this domain.

References

- Alibali, Martha W., Dana C. Heath, & Heather J. Myers (2001). Effects of visibility between speaker and listener on gesture production: Some gestures are meant to be seen. *Journal of Memory and Language*, *44*, 169-188.
- Bavelas, Janet B., & Nicole Chovil (2000). Visible acts of meaning. An integrated message model of language use in face-to-face dialogue. *Journal of Language and Social Psychology*, *19*, 163-194.
- Bavelas, Janet B., Nicole Chovil, Linda Coates, Lori Roe (1995). Gestures specialised for dialogue. *Personality & Social Psychology Bulletin*, *21*, 394-405.
- Bavelas, Janet B., Nicole Chovil, Douglas A. Lawrie, & Allan Wade (1992). Interactive gestures. *Discourse Processes*, *15*, 469-489.
- Bavelas, Janet B., Jennifer Gerwing, Chantelle Sutton, & Danielle Prevost (2008). Gesturing on the telephone: Independent effects of dialogue and visibility. *Journal of Memory and Language*, *58*, 495-520.
- Bavelas, Janet B., Christine Kenwood, Trudy Johnson, & Bruce Phillips (2002). An experimental study of when and how speakers use gestures to communicate. *Gesture*, *2*, 1-17.
- Butterworth, Brian L. (1975). Hesitation and semantic planning in speech. *Journal of Psycholinguistic Research*, *4*, 75-87.
- Chafe, Wallace L. (1976). Givenness, contrastiveness, definiteness, subjects, topics, and point of view. In Charles N. Li (Ed.), *Subject and topic* (pp. 25-55). New York: Academic Press.

- Clark, Herbert H. (1996). *Using language*. Cambridge: Cambridge University Press.
- Clark, Herbert H., & Catherine R. Marshall (1981). Definite reference and mutual knowledge. In A. K. Joshi, B. Webber, & I. Sag (Eds.), *Elements of discourse understanding* (pp. 10-63). Cambridge: Cambridge University Press.
- Clark, Herbert H., & Donna Wilkes-Gibbs (1986). Referring as a collaborative process. *Cognition*, 22, 1-39.
- Fetzer, Anita, & Kerstin Fischer (2007) (Eds). *Lexical markers of common grounds*. London: Elsevier.
- Foraker, Stephanie, & Susan Goldin-Meadow (2007). *Gesture and discourse: How we use our hands to refer back*. Talk presented at the 3rd International Society for Gesture Studies Conference: Integrating Gestures, June 18-22, Chicago, IL.
- Furuyama, Nobuhiro (2000). Interaction between instructor and learner. In David McNeill (Ed.), *Language and gesture*. Cambridge: Cambridge University Press.
- Fussell, Susan R., & Robert M. Krauss (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.
- Gerwing, Jennifer, & Janet B. Bavelas (2004). Linguistic influences on gesture's form. *Gesture*, 4, 157-195.
- Gullberg, Marianne (2006). Handling discourse: Gestures, reference tracking, and communication strategies in early L2. *Language Learning*, 56, 155-196.
- Halliday, Michael A. K. (1967). Notes on transitivity and theme in English. *Journal of Linguistics*, 3, 199-244.
- Haviland, Susan E., & Clark, Herbert H. (1974). What's new? Acquiring new information

as a process in comprehension. *Journal of Verbal Learning and Verbal Behavior*, *13*, 512-521.

Holler, Judith (2003). *Semantic and pragmatic aspects of representational gestures: towards a more unified model of communication*. Unpublished doctoral dissertation, University of Manchester, UK.

Holler, Judith, & Geoffrey Beattie (2002). A micro-analytic investigation of how iconic gestures and speech represent core semantic features in talk. *Semiotica*, *142*, 31-69.

Holler, Judith, & Geoffrey Beattie (2003a). How iconic gestures and speech interact in the representation of meaning: are both aspects really integral to the process? *Semiotica*, *114*, 86-116.

Holler, Judith, & Geoffrey Beattie (2003b). Pragmatic aspects of representational gestures: Do speakers use them to clarify verbal ambiguity for the listener? *Gesture*, *3*, 127-154.

Holler, Judith, & Rachel Stevens (2007). An experimental investigation into the effect of common ground on how speakers use gesture and speech to represent size information in referential communication. *Journal of Language and Social Psychology*, *26*, 4-27.

Holler, Judith, & Katie Wilkin (2009). Communicating common ground: how mutually shared knowledge influences the representation of semantic information in speech and gesture in a narrative task. *Language and Cognitive Processes*, *24*, 267-289.

Hostetter, Autumn, Martha Alibali, & Sotaro Kita (2007). I see it in my hands' eye: Representational gestures reflect conceptual demands. *Language and Cognitive Processes*, *22*, 313-336.

- Isaacs, Ellen A., & Herbert H. Clark (1987). References in conversations between experts and novices. *Journal of Experimental Psychology: General*, 116, 26-37.
- Jacobs, Naomi, & Alan Garnham (2007). The role of conversational hand gestures in a narrative task. *Journal of Memory and Language*, 56, 291-303.
- Kendon, Adam (2004). *Gesture: Visible action as utterance*. Cambridge: Cambridge University Press.
- Keysar, Boaz, Dale J. Barr, Jennifer A. Balin, & Timothy S. Paek (1998). Definite reference and mutual knowledge: Process models of common ground in comprehension. *Journal of Memory and Language*, 39, 1-20.
- Kuno, Susumu (1972). Functional sentence perspective. *Linguistic Inquiry*, 3, 269-320.
- Levy, Elena, & David McNeill (1992). Speech, gesture, and discourse. *Discourse Processes*, 15, 277-301.
- McNeill, David (1992). *Hand and mind: What gestures reveal about thought*. Chicago: University of Chicago Press.
- McNeill, David (2005). *Gesture and thought*. Chicago: University of Chicago Press.
- McNeill, David, Justine Cassell, & Elena Levy (1993). Abstract deixis. *Semiotica*, 95, 5-20.
- Oldfield, R. Carolus (1971). The assessment and analysis of handedness: The Edinburgh inventory. *Neuropsychologia*, 9, 97-113.
- Özyürek, Asli (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.

- Parrill, Fey (2010). The hands are part of the package: gesture, common ground and information packaging. In John Newman, & Sally Rice (Eds.), *Empirical and experimental methods in cognitive/functional research*. Stanford: CSLI.
- Pine, Karen J., Hannah Bird, & Elizabeth Kirk (2007). The effects of prohibiting gestures on children's lexical retrieval ability. *Developmental Science*, 10, 747-754.
- Prince, Ellen (1981). Toward a taxonomy of given-new information, In Peter Cole (Ed.), *Radical pragmatics* (pp. 223–255). New York: Academic Press.
- Sacks, Harvey, Emanuel A. Schegloff, & Gail Jefferson (1974). A simplest systematics for the organization of turn-taking in conversation. *Language*, 50, 696-735.

Acknowledgements

We would like to thank Gale Stam and Mika Ishino, as well as independent reviewers, for their valuable comments on an earlier draft of this manuscript. We would also like to thank the Economic and Social Research Council for funding this research (research grant RES-061-23-0135) as well as the participants taking part in our study.

Tables and Figures

Table 1. Overview of average proportions of references classed as definite or indefinite accompanied by gesture or no gesture (in total as well as for individual gesture categories).

Accompaniment	Reference type	
	Definite (the/that)	Indefinite (a/this)
No gesture	0.29	0.14
Gesture (all categories combined)	0.71	0.86
Split up by category:		
Iconic	0.86	0.75
Deictic	0.06	0.15
Metaphoric	0.01	0.00
Beats	0.03	0.04
Interactive	0.05	0.06

Table 2. Average proportions (and frequencies) of definite and indefinite references accompanied by ‘action’ and ‘entity’ gestures..

Reference type	Action	Entity	Total
Definite (‘the’/‘that’)	62.71% (74)	37.29% (44)	100% (118)
Indefinite (‘a’/‘this’)	37.84% (28)	62.16% (46)	100% (74)

Table 3. Average proportions (and frequencies) of ‘action’ and ‘entity’ gestures used in the two experimental common ground conditions (common ground and no common ground).

Condition	Action	Entity	Total
CG	60.9% (53)	39.1% (34)	100% (87)
NCG	46.7% (49)	53.3% (56)	100% (105)

Figure 1. Overview of the mean percentage of ‘action’ and ‘entity’ gestures accompanying definite and indefinite references in the two experimental common ground conditions.

