

The Interaction of Iconic Gesture and Speech in Talk

Judith Holler and Geoffrey Beattie

University of Manchester
United Kingdom

Abstract. One traditional view of how speech and gesture interact in talk is that gestures represent information, which is largely redundant with respect to the information contained in the speech that they accompany. Other researchers, however, have primarily stressed a complementary interaction of gesture and speech, and yet others have emphasised that gesture and speech interact in a very flexible manner. These discrepant views have crucially different implications with regard to the communicative role of gestures. The study reported here offers a systematic and detailed investigation of this issue to gain further insights into how the two modalities interact in the representation of meaning. The findings support the notion of gesture and speech interacting in a highly flexible manner.

1 Introduction

Iconic gestures are spontaneous movements of the hands and arms that accompany talk, and they represent concrete meaning that is closely associated with the semantic information in the speech. The way in which these hand gestures interact with speech in the representation of meaning has been described in the literature in quite different ways. Whereas one view of the gesture-speech relationship is that gestures convey mainly information that is redundant with respect to the speech (e.g., [1], [2]), others have stressed their complementary relationship with speech (e.g., [3]). Yet other investigators have been rather vague about the exact relationship of iconic gesture and speech by saying that these gestures ‘illustrate’ what is being said (e.g., [4], [5]) – which can be understood in a variety of ways, for example in terms of the gestural information emphasising what is said (i.e., representing the same information) or it could be understood as meaning that the gestural information complements the verbal information (i.e., by representing additional semantic information).

An example of a gesture representing information which can be considered as largely redundant with regard to the speech, is the following (example 1):

(1)

‘a group of..of men, kind of around a [big table]’¹

[both hands rise to about stomach height, they are held in parallel at the right and the left side of the body, the palms are facing towards the middle]

(Extracted from authors’ corpus)

¹ For transcription conventions, please see Appendix.

In this example, the gesture refers to the table that the speaker is talking about and it emphasises its size. Thus, the gesture can be considered as not providing any information over and above the speech, which also contains the information that the table is big. This gesture would therefore correspond to Birdwhistell's [1] and Krauss, Morrel-Samuels and Colasante's [2] view that the information represented by gesture and speech is largely redundant. Although Birdwhistell [1] did not explicitly refer to iconic gestures, he did include in his analyses hand gestures that are essentially iconic in that they represent concrete semantic information. For example, Birdwhistell referred to gestures that show how an action is being carried out (e.g., 'swiftly' or 'slowly'), which he referred to as kinesic markers of the 'manner of action'. Krauss *et al.* ([2], p.743) also stated that 'although gestures can convey some information, they are not richly informative, and the information they convey is largely redundant with speech'. (Although Birdwhistell and Krauss *et al.* arrived at a similar conclusion concerning the semantic interaction of gesture and speech, it has to be noted that their views of the communicative role of gestures differ crucially.)

McNeill [3] seemed to stress a rather strict complementary pattern of gesture-speech interaction. In arguing for a broader view of language that takes into account gesture McNeill explained that gesture and speech often provide information about *different* semantic aspects of the same scene to make clear that a crucial part of a speaker's message might be ignored if only the speech is considered as conveying semantic information. The following example represents a case of speech and iconic gesture being complementary in their representations (example 2):

(2)

'she [chases him out again]'

[hand appears to swing an object
through the air]

(Extracted from [3], p.13)

Whereas the speech provides here the information that one character is chasing another, the accompanying iconic gesture shows that an instrument (an umbrella) is being used to do so. Although speech and gesture have here the same semantic reference point as they both represent information relating to the same event, they provide information about different aspects of this event so that they interact in a very complementary fashion. Hence, McNeill [3] argued that speech and gesture together reveal a fuller insight into a speaker's thoughts.

Rather than emphasising either a redundant or a complementary pattern of gesture-speech interaction, Kendon (e.g., [6], [7]) argued for a very flexible interaction of gesture and speech, with gesture representing an 'available resource' that speakers can use to respond to very different communicational demands.

Overall, it appears that the views of how speech and gesture predominantly interact are somewhat discrepant, and it thus seems necessary to investigate this issue further to gain more of an idea how these two channels of communication interact in the representation of meaning. However, instead of providing a number of individual examples, which show that speech and gesture can interact in one way or the other, as many previous investigations into this issue have done, this study systematically analyses a larger corpus of gestures.

2 Experimental Investigations into the Semantic Interaction of Gesture and Speech

2.1 An Attempt to Quantify the Interaction of Iconic Gesture and Speech

An experimental investigation was carried out investigating the interaction of gesture and speech by using a detailed analysis of a corpus of gesture-speech compounds produced in the context of narratives. These narratives stemmed from speakers telling cartoon stories, which were projected onto a wall in front of them, to the experimenter (JH). The gestural and verbal utterances referred to six ‘semantic events’ that formed part of two different Tom and Jerry cartoon stories (for example, ‘Jerry stabbing Tom in the tail with a sharp instrument’, ‘Spike, the dog, dangling Jerry by his tail’, etc.). For the analysis, 58 verbal utterances accompanied by 58 iconic gestures (stemming from 27 different speakers) were selected and the semantic information represented by gesture and speech was scored according to 20 detailed semantic categories (the general idea of applying a semantic feature approach was derived from Beattie and Shovelton [e.g., 8, 9, 10, 11]. The semantic categories used referred to, for example, information about the kinds of entities involved in the event that the speakers talked about (such as ‘agent’, ‘object [in terms of an entity that is being acted upon] and ‘instrument’), the direction of, or the force associated with, a movement, the position of individual entities relative to each other and the surrounding space, as well as the shape and the size of the individual entities that formed part of the event (see Table 1; for more detail on the individual semantic categories, please see [12]). An important aspect of the study is also that it took into account information that was only implicitly represented in speech, as well as explicitly represented information. The scoring scheme applied in the study was designed to capture such subtle differences.

The mathematical scoring scheme was developed to quantify the information represented by gesture and speech. It consisted of certain ‘informational values’, namely 0 (meaning that information was not represented), 0.5 (meaning that information was *implicitly* represented) and 1 (meaning that information was *explicitly* represented). Scores were given for all 58 gestures and 58 speech extracts in separate, and each gesture and speech extract was scored according to the 20 semantic categories displayed in Table 1 (for more detail concerning the criteria used for scoring the information, please see [12]). The inter-observer reliability for scoring the information represented by the gestures and the speech extracts was calculated using Cohen’s Kappa, which resulted in $K=.96$ for speech and $K=.97$ for gesture.

The informational value 0.5 did not find appliance in the scoring of the gestural information. Hence, the possible interaction patterns of gesture and speech could be categorised into six different types, as displayed in Table 2.

The next step in the analysis involved calculating how often gesture and speech had interacted in the six different ways identified here concerning each of the semantic categories. In order to simplify the rather complex database resulting from this procedure (i.e., a 6x20 matrix), the predominant interaction pattern was identified for each semantic category. Those semantic features with clear peaks concerning one of the interaction patterns are shown in Table 3.

Table 1. Overview of the basic and their related, more detailed semantic categories applied

| | |
|--------------------------|---|
| Entity | 1. Agent 2. Object 3. Instrument |
| Action | 4. Body-parts involved in the movement 5. Direction of the movement 6. Point of contact defined 7. Force |
| Relative position | 8. Relative position: Agent – Object 9. Relative position: Agent – Instrument 10. Relative position: Object – Instrument 11. Relative position: Object – Space |
| Size | 12. Size: agent 13. Size: object 14. Size: instrument |
| Shape | 15. Agent 16. Object 17. Instrument |
| Shape of a part | 18. Agent 19. Object 20. Instrument |

(From [12])

Table 2. Overview of the six types of speech-gesture interaction patterns resulting from the scoring scheme applied

| Speech - Gesture | |
|------------------|---|
| 0-0 | The information is represented <i>neither by speech nor by gesture</i> . |
| 0-1 | The information is represented <i>only gesturally</i> . |
| 1-0 | The information is represented <i>only verbally</i> . |
| 0.5-0 | The information is <i>implicitly</i> represented in <i>speech, but not in gesture</i> . |
| 0.5-1 | The information is <i>implicitly</i> represented in <i>speech and explicitly in gesture</i> . |
| 1-1 | The information is <i>explicitly</i> represented in <i>both gesture and speech</i> . |

(From [12])

Complexity and Flexibility – Two Important Aspects Characterising the Interaction of Iconic Gesture and Speech.

What can be seen from Table 3 is, first of all, that the interaction of gesture and speech seems rather complex, as gesture and speech were shown to interact in at least five different ways, and these include that gesture and speech represent both complementary as well as redundant information. The analysis also reveals that it is necessary to be more precise when we talk about ‘complementarity’ in association with the semantic interaction of gesture and speech. This is because the gestural and the verbal information can be strictly complementary in

Table 3. Overview of the semantic features that were predominantly represented by one of the six gesture-speech interaction patterns identified

| Speech - Gesture | Semantic feature |
|------------------|---|
| 0-0 | Action: force Rel. position: O-S Size: instrument |
| 0-1 | Rel. position: A-O Rel. Position: A-I Shape of a part: object |
| 1-0 | ---- |
| 0.5-0 | Size: object Shape: object |
| 0.5-1 | Action: body-parts Shape of a part: agent |
| 1-1 | Entity: agent Entity: object Entity: instrument |

(From [12])

that each of the two communicational channels represents information that is not at all represented in the respective other (for example, as in the case of the gesture-speech example from McNeill described above). However, they can also complement in that speech represents some (i.e., implicit) information, while the gesture represents the same information more explicitly. Furthermore, although some of the semantic features were shown to be predominantly represented by gesture and speech interacting in one certain way, each individual semantic feature was represented by iconic gesture and speech interacting in a variety of different ways.

Overall, it seems that stressing either a redundant or a complementary interaction of iconic gesture and speech does not provide an accurate picture of how the two modalities actually interact in the representation of meaning. Rather, it seems necessary to stress the complexity that characterises the interaction of iconic gesture and speech. Furthermore, it is important to note that iconic gesture and speech can interact in a very flexible manner, even concerning the representation of the same semantic aspects. Hence, the findings of this quantitative analysis support Kendon's (e.g., [6], [7]) notion of how speech and gesture interact.

Having come to this conclusion, the question remains as to what factors do then have an impact on how gesture and speech interact. One such factor could of course be the social context in which spoken discourse is embedded, and in line with Kendon, it could be assumed that this context is directly involved in shaping the communicational demands that a speaker responds to using gesture.

Based on the observation that in some cases in which speech provided only implicit information, this information was explicitly represented by gesture, one could assume that one communicative function of gesture may be the facilitation of inferences that listeners have to make when information is only implicitly provided. In other words, the gestures in these cases might provide the recipient with semantic

cues as to what the correct inferences are. If future investigations could provide evidence that hand gestures really do function as an aid for the recipient to draw inferences during talk, this would be evidence for a very important communicative function of gesture, because inferences play such a crucial role in the understanding of everyday talk. Moreover, it would show that gestures are indeed used by speakers to facilitate the listener's understanding and thus that the use of gesture is influenced by the social context. Hence, the findings of the present study suggest a potential pragmatic use of gesture, which calls for further investigation, particularly considering that there is still much debate about whether gestures are communicatively intended or not (cf. [13], [14]).

3 Conclusions

Overall, the investigation reported here shows that the interaction of iconic gesture and speech is complex as well as flexible. It thus clarifies the issue of how iconic gesture and speech interact in the representation of meaning somewhat by showing that views emphasising either a predominantly complementary or a predominantly redundant pattern of interaction do not reflect the most crucial aspect that seems to characterise the interaction of iconic gesture and speech, namely flexibility. Furthermore, the findings support the notion that speakers draw on gesture in order to fulfil certain communicational functions. Because the scoring scheme applied in this study took into account inferences that listeners might have to make in order to comprehend spoken discourse, it was possible to identify an important potential use of gesture, namely the facilitation of inferences. Future research is needed to further test whether gestures are indeed used by speakers for this purpose.

References

1. Birdwhistell, R. L.: *Kinesics and Context. Essays on Body-Motion Communication*. Allen Lane The Penguin Press, London (1970).
2. Krauss, R. M., Morrel-Samuels, P., Colasante, C.: Do conversational hand gestures communicate? *Journal of Personality and Social Psychology* (1991), 61, 743-754.
3. McNeill, D.: *Hand and Mind: What Gestures Reveal about Thought*. University of Chicago Press, Chicago (1992).
4. Ekman, P., Friesen, W. V.: The Repertoire of Nonverbal Behaviour: Categories, Origins, Usage and Coding. *Semiotica* (1969), 1, 49-98.
5. Argyle, M.: *Bodily Communication*. Methuen and Co Ltd., London (1975).
6. Kendon, A.: Gesticulation and speech: Two aspects of the process of utterance. In: Key, M. R. (ed.). *The Relationship of Verbal and Nonverbal Communication*. Mouton, The Hague (1980), 207-227.
7. Kendon, A.: Some Uses of Gesture. In: Tannen, D. and Saville-Troike, M. (eds.), *Perspectives on Silence*. Ablex Publishing Corporation, Norwood New Jersey (1985) 215-234.
8. Beattie, G., Shovelton, H.: Do Iconic Hand Gestures Really Contribute Anything to the Semantic Information Conveyed by Speech? An Experimental Investigation. *Semiotica* (1999), 123, 1-30.
9. Beattie, G., Shovelton, H.: Mapping the Range of Information Contained in the Iconic Hand Gestures that Accompany Spontaneous Speech. *Journal of Language and Social Psychology* (1999), 18, 438-462.

10. Beattie, G., Shovelton, H.: An experimental investigation of the role of different types of iconic gesture in communication: A semantic feature approach. *Gesture* (2001), 1, 129-149.
11. Beattie, G., Shovelton, H.: An Experimental Investigation of Some Properties of Individual Iconic Gestures that Mediate their Communicative Power. *British Journal of Psychology* (2002), 93, 179-192.
12. Holler, J., Beattie, G.: How Iconic Gestures and Speech Interact in the Representation of Meaning: Are Both Aspects Really Integral to the Process? *Semiotica* (2003), 146, 81-116.
13. Butterworth, B., Hadar, U.: Gesture, speech, and computational stages: A reply to McNeill. *Psychological Review* (1989), 96, 168-174.
14. Krauss, R. M., Chen, Y., Gottesman, R. F.: Lexical gestures and lexical retrieval: A process model. In: McNeill, D. (ed.), *Language and Gesture*. Cambridge University Press: Cambridge (2000), 261-283.

Appendix

Transcription conventions:

Segments of speech analysed are marked using 'single quotes'. That part of the verbal utterance that was accompanied by the iconic gesture is marked using [square brackets]. The iconic gesture that accompanied the verbal utterance is described underneath the extract of speech and this description is also contained within square brackets.