

Jan Peter de Ruiter

THE FUNCTION OF HAND GESTURE IN SPOKEN CONVERSATION

ABSTRACT

In this paper, I will discuss the ongoing controversy in the gesture literature about the function of hand gesture in spoken conversation.¹ Two ›camps‹ can be identified: those who believe that gesture enhances communication by providing extra information to the listener and, on the other hand, those who believe that gesture is not communicative, but rather that speakers gesture to facilitate their own word-finding processes. I will review a number of important studies regarding this controversy, and conclude that the empirical studies that are cited to support the claim that gesturing is done for speaker-internal reasons are in fact supporting the notion that gesture is a communicative device which is *compensating* for problems in speech by providing information in another channel.

1. TYPES OF GESTURE

When people speak, their speech is often accompanied by spontaneous hand gestures. There are several typologies around that distinguish between different classes of gestures. In De Ruiter (2000) I have presented a typology that is based on McNeill (1992), but is adapted for use in cognitive psychological modeling by emphasizing the underlying mental *representations* of gesture, rather than the semiotic relation with the concurrent speech.² This typology is given below.

1. Emblems. These gesture have a form-meaning mapping that is a) arbitrary (idiosyncratic) and b) lexicalised. The meaning of emblems is therefore language dependent. For example: a thumbs-up gesture in German means ›OK‹, while in some Mediterranean countries it has a very obscene meaning, comparable to the meaning of the ›middle finger‹ in German.
2. Iconic gestures. The shape of iconic gestures resembles the shape of their referent, but there are no lexicalised conventions for these representations. In other words, speakers who are employing iconic gestures are not bound by linguistic conventions in how they choose to display a certain concept in gesture – the way they display it is spontaneous. In this typology, the category that McNeill calls ›metaphoric‹ gestures is also considered to be iconic, the reasoning being

that for the underlying gestural representation it does not matter whether the referent of the gesture is real or abstract.

3. Pantomimic gestures. These gestures are ›enactments‹ of motoric actions. For instance, performing the motion that corresponds to throwing a ball without actually holding a real ball.
4. Deictic gestures. These are conventionally known as ›pointing gestures‹. They indicate locations or directions. A large subset of deictic gestures is *linguistically obligatory*, in the sense that the accompanying speech makes no sense without the gesture. For instance, when I say ›please put your coat over there‹, without indicating with a deictic gesture where ›there‹ is, the statement does not make sense.
5. Beats. Beats are rhythmic motions of the hand or fingers that seem not to bear any representational meaning, like iconic or pantomimic gestures do. Beats have been claimed to be related to the phonology of the accompanying speech (McNeill), although McClave (1994) found only very little evidence supporting this claim.³

1.1. THE COMMUNICATIVE FUNCTION OF GESTURES

For some of the types of gesture mentioned in the typology above, it is clear and uncontroversial that they are intended to communicate information for the benefit of the listener. Emblems, for instance, are often used to *replace* speech. Instead of saying ›OK‹ one can make the OK gesture. Pointing gestures can also be assumed to be communicative, for in many cases the speech does not even make sense without them. But even for pointing gestures or emblems that are strictly speaking redundant (with respect to the accompanying speech) there seems to be a consensus in the gesture literature that they constitute a communicative phenomenon.

This is different for the other categories. Beats have been claimed by McNeill to have a meta-narrative function, reflecting the structure of the discourse. The evidence McNeill presents for this claim is of a qualitative and anecdotal nature, so further studies involving a larger number of gestures are needed to substantiate this claim. About pantomimic gestures, hardly any claims exist in the literature, probably because the category is not large, relatively understudied, and not distinguished as such by many authors. The question whether iconic gestures have a communicative function, however, is a hotly debated issue.

In the remainder of this paper, I want to give the reader an overview about the different points of view that exist regarding this issue, and summarize some of

the empirical evidence for and against the communicative function of iconic gesture. Also, I will try to show that all experimental results discussed here are compatible with the hypotheses that a) gesture is a communicative device, and b) gesture and speech are mutually compensating for difficulties in the other channel. In the remainder of this paper, the word »gesture« is intended to mean »iconic gesture« unless indicated otherwise.

1.2. THE COMMUNICATIVE VIEW

The most outspoken proponent of the view that gesture is a communicative device is Adam Kendon.⁴ Kendon holds that a careful study of gesture should lead one to the conclusion that gestures are made for the benefit of the listener. The fact that people indeed frequently gesture on the telephone, an argument often used by opponents of the communicative view, is explained by Kendon by assuming that people are so used to gesturing during speaking that they are unable to suppress it on the telephone.⁵ In De Ruiter (2000), the claim is made not only that gesture is a communicative device, but also that gesture and speech are *mutually compensatory channels*: if there is trouble expressing something in speech, gestures are employed to provide the information to the listener in another channel. If, on the other hand, gesturing is difficult or not possible, speech will have to become more elaborate to compensate for the lack of the gesture channel.⁶ In this paper, I will argue that this hypothesized mechanism, henceforth called the *channel compensation mechanism* (CMM), can adequately and parsimoniously explain all the experimental findings brought forth by the adversaries of the communicative view.

There is to date little experimental work that directly addresses the question whether gesture information is perceived and integrated by the listener. Cassell, McNeill & McCullough (1999) had an accomplice tell a story with either gestures that matched the speech or did not match it (i. e. conflicted with the information in the speech). They showed that in both cases, listeners do pick up on the information provided in the gesture channel.⁷

1.3. THE NON-COMMUNICATIVE VIEW

The view that gestures are not communicative is defended in – among others – Rimé & Schiaratura (1991), Krauss, Morrel-Samuel, & Colasante (1991), Krauss, Chen & Gottesmann (2000), and Hadar & Butterworth (1997).⁸ In these publica-

tions, the general claim is that (iconic, or as Krauss prefers to call them, »lexical«) gestures are hardly interpretable without the concurrent speech. Gestures are therefore not assumed to serve the comprehension of the listener, but to facilitate processing in the speaker him- or herself. In Krauss et al. (2000) this notion is made more explicit in the formulation of an outline of a computational model.⁹

1.4. THE FACILITATION HYPOTHESIS: SOME THEORETICAL NOTES

Although several authors claim that making gestures facilitates the word-finding process during speech, they are often not very clear about *how* this is supposed to happen. Performing an iconic gesture involves a certain hand motion of which the shape and the trajectory through space share some of the characteristics of the referent. How doing so can facilitate the process of lexical retrieval (finding the correct word to produce) is not self-evident, and therefore in urgent need of explanation.

The most explicit and detailed account of how this facilitation could work is in Krauss, Chen & Gottesmann (2000), in which a rough outline of a computational model is given.¹⁰ Facilitation of lexical retrieval is a process that is described by the authors as a form of »cross-modal priming« by the output of the (kinesic) self-monitoring of the gesture. In other words, the proprioception and interpretation of one's own motor patterns during gesturing provides system-internal information that assists the process of word-finding.¹¹ There are a priori – that is, before even looking at empirical data – a number of problems with this proposal.

First, in order to be able to generate an iconic gesture, spatial properties of the concept to be gestured about must be translated into manual behavior. This means that the information that is to facilitate the lexical retrieval is already available *before* a gesture is performed; this information is needed in order to generate the correct motor program for the gesture. Why would our speech production system first encode this information into gesture – an external representation visible to the outside world – and then decode the spatial features from its own gesture via kinesthetic self-awareness back into visual features when it already had access to the spatial features? It is rather implausible that a lexical retrieval facilitation system has evolved that is so inefficient, inelegant, and redundant.

A second problem stems from the frequently made observation that iconic gestures often do not correspond or refer to a single word in the speech.¹² Frequently, the so-called »speech affiliate« of the gesture is a phrase consisting of several words, and frequently is not even *present* in the speech. An example of the

latter (from De Ruiter, 2000) is a fragment in which a speaker says (translated from Dutch) »and then one can see a *bird watchers society*«. While uttering the italicized words of that phrase, the speaker used both hands to outline a rectangular shape in front of his body. Only for those who have seen the cartoon that the subject is describing, it is clear that the gesture refers to a big sign appearing in the cartoon which says »bird watchers' society«. But there is *nothing* in the speech that refers to this sign. It is therefore unlikely that the gesture in this case is made to assist in lexical retrieval, and it is more plausible to assume that the gesture is performed for communicative purposes.

II. THE FACILITATION HYPOTHESIS: EMPIRICAL EVIDENCE

II.1. WORD FAMILIARITY AND GESTURE TIMING

Morrell-Samuels & Krauss (1992) have found that the onset of gestures usually precedes the word they are affiliated to.¹³ Furthermore, the less familiar (frequent) a word is, the larger is the time interval by which the gesture precedes the speech. The authors claim that this makes it at least plausible that gestures facilitate lexical retrieval, because gesture has more time to »help« when the less familiar words are produced. At least as plausible is the assumption that gesture and speech are relatively independent processes that start at about the same point in time. Familiar words are known to be produced faster and will therefore occur earlier in time, relative to the gesture.¹⁴

II.2. THE INTERPRETABILITY OF GESTURES

Krauss, Morrell-Samuels & Colasante (1991) argue that since gestures are hardly interpretable in isolation (i. e. without the concurrent speech), they cannot have a communicative function.¹⁵ This argument is flawed. In order for gesture to be communicative, gestures do not need to make sense to listeners in isolation, because listeners *do* have access to the accompanying speech, and as the work of McNeill (1992) and other has repeatedly shown, gestures and their accompanying speech form a tight semantic unit, sometimes redundant, but mostly complementing each other. This means that in the normal situation, namely that the listener has access to gesture and speech at the same time, the interpretation of the gesture is not hard at all, and can therefore perform an important role in the communication.

II.3. GESTURING ON THE TELEPHONE

Another argument often used by proponents of the non-communicative view is that people frequently gesture while they are on the telephone or, more generally, when there is no visual contact between speaker and listener. This phenomenon could be adequately explained by the theory that gestures facilitate the speaking process. If people gesture for speaker-internal reasons, they will do it on the telephone as well.

However, it is known that speakers gesture more when they have visual contact with the listener.¹⁶ Also, the fact that people gesture on the telephone does not rule out the possibility that gestures are nevertheless communicatively intended. Gesturing could be so intricately linked to speaking that it is hard to suppress gesturing when speaking on the telephone. After all, speaking with someone whom you cannot see is, from the viewpoint of human evolution, a very recent invention. It is conceivable that if gesturing is deeply integrated with the speaking process – as for instance McNeill (1992) assumes – the mere fact that the addressee is invisible is not sufficient to cause people to stop or suppress gesturing.

II.4. PREVENTING GESTURE

The most direct way of testing the hypothesis that gesturing facilitates the speaking process itself would be to prevent participants from gesturing and look at the effects it has on their speech.

Graham & Argyle (1975) presented geometrical line drawings to what they called »encoders«.¹⁷ Encoders were either native speakers of Italian or native speakers of English. The task of the encoder was to describe those drawings to a »decoder« who had to reproduce the drawing. In one condition the encoder was allowed to gesture, while in the other he or she was not. The accuracy of the reproduction was higher when the encoder was allowed to gesture. This effect was even stronger for those drawings that were rated to be of low codability, demonstrating that the information presented in the encoder's gesture had a positive effect on the communication between encoder and decoder. No effects on the content of the speech were found. In Graham & Heywood (1975) essentially the same experiment was performed with only English speaking participants.¹⁸ They coded a large number of speech related dependent variables, of which only a few turned out to differ between the gesture and the no-gesture condition. Specifically, the elimination of gestures led to an increase in expressions describing spatial re-

lations and to a decrease in the number of demonstratives. Also, the time spent pausing (in speech) increased in the no-gesture condition. As the authors note, these findings need not be explained by the assumption that the production of speech is facilitated by gesturing. Rather, it is likely that the increased number of phrases describing spatial relations and the increased pausing time are a compensation for not being able to use the gesture modality, as is also suggested by Kendon (1972) and by De Ruiter (2000, the previously mentioned CCM hypothesis).

Rimé, Schiaratura, & Ghysseleux (1984) let their participants engage in free conversation about certain predefined general topics.¹⁹ During the second half of the conversation, the head, hand and arm movements of the participant were immobilized by devices attached to the armchair of the participant. It was found that the vividness of the imagery in the speech *decreased* when the hands were immobilized. At first sight these results seem to contradict the aforementioned findings by Graham & Argyle (1975) and Graham & Heywood (1975): they found an *increase* in »spatial« speech, while Rimé et al. found a *decrease* in spatial speech.²⁰ However, a crucial difference is that in the studies by Graham & Heywood and Graham & Argyle, the participants were requested to speak about the presented line drawings, while in the study by Rimé et al. participants were free to select the content of their speech. Assuming, again, that gesture is a communicative device that serves especially well to transmit spatial information, and following the logic of the CCM hypothesis, in the studies by Graham & Argyle and Graham & Heywood participants were forced to compensate for the lack of gesture by producing more spatial descriptions in speech, while in the study by Rimé et al. participants could avoid talking about topics containing spatial information, thereby circumventing the problems the participants of the Graham & Argyle and the Graham & Heywood studies had.

Finally, Rauscher, Krauss & Chen (1996) prevented their participants from gesturing as well.²¹ The participants in their study had to describe cartoon animations to listeners, while during half of the time they were not allowed to move their hands. Their findings were (1) that speech with spatial content was less fluent when gesturing was not permitted, (2) speech without spatial content was not affected, and (3) that the frequency of (non-juncture) filled pauses in the speech increased in the no-gesture condition, but only when the participants were producing speech with spatial content. The authors conclude from these findings that gesture facilitates access to the mental lexicon, for the effects of preventing gesture are similar to those of word-finding difficulties. However, these results can easily and in fact more parsimoniously be interpreted as evidence that gesture functions as a compensatory communicative device, as in the CCM hypothesis

mentioned previously. Given that the gesture modality is much more efficient in expressing spatial information, the loss of fluency in the no-gesture condition is predictable: the generation of speech with spatial content needs to be adapted (i. e. be more accurate and elaborate) when the gesture modality is unavailable. If the content of the speech is not spatial, this problem does not occur, which is exactly what the authors found. The authors' conclusion that their findings indicate that gesturing facilitates lexical access therefore seems unwarranted.

II.5. A DIRECT TEST: THE TIP OF THE TONGUE STUDY

Beattie & Coughlan (1999) provided the most direct test of the hypothesis that gesture facilitates lexical retrieval.²² They elicited Tip Of the Tongue (TOT) states in subjects. TOT states are states in which we are sure that the information [about the word we are looking for] is in memory but are temporarily unable to access it.²³ Beattie and Coughlan allowed half of their subjects to gesture, but asked the other half to keep their arms folded during the experiments, effectively preventing them from gesturing. Then the subjects were presented with definitions and were then asked which word is described by the provided definition.

The prediction of the lexical retrieval facilitation hypothesis in this experiment is that the people who are allowed to gesture will »resolve« (in the sense of »finally finding the word they are looking for«) a higher proportion of TOT states than those who had their arms folded. This turned out not to be the case. The main finding of this study was that »significantly more TOT states were resolved when gestures were absent than when they were present«.²⁴ In other words, the presence of gesture *lowered* the probability of resolving the TOT.

This finding is very problematic for the proponents of the view that gestures facilitate lexical retrieval.

The CCM hypothesis can actually explain this – at first sight paradoxical – finding that gesturing lowers the probability of resolving the word finding difficulty: the more difficult it is to find the word, the more the gesture channel will be activated to compensate for the impending »trouble« in the speech channel. In other words, the presence of gesture indicates that the word-finding difficulties are more serious and therefore less likely to be resolved.

III. CONCLUSIONS

In comparing the evidence in favor of and against the communicative hypothesis and its rival, the facilitation hypothesis, the following balance can be made.

1. Although it seems plausible that gesture is a communicative device, there is little direct evidence in favor of the view that gesture actually does transmit essential information to the listener.
2. Most of the studies done with the »preventing gesture« paradigm support the notion that gesture and speech are mutually compensating systems, as suggested in De Ruiter (2000). This even holds for those studies that were claimed to support the view that gesture does not communicate but rather serves to facilitate word retrieval processes.
3. The only study that provides a direct empirical test of the lexical retrieval facilitation hypothesis, namely the TOT study by Beattie & Coughlan, yields results that are clearly falsifying this hypothesis.

Although the discussion will probably continue in the literature, it seems that the communicative view is royally ahead on points, while the facilitation hypothesis has very little evidence to support it. Partly this is due to the fact that in every study that reveals that problems in speech lead to more gestures, the Compensatory Channel Mechanism can always explain this rather straightforwardly. To rule out this alternative explanation, more studies like the Beattie & Coughlan (1999) study that explores TOT states in speakers, or the Cassell et al. (1999) study that focuses on the information uptake by the listener, are needed to resolve this issue.

1 The author can now be contacted at: janpeter.deruiter@mpi.nl

2 Jan Peter de Ruiter: The production of gesture and speech, in: David McNeill (Ed.): *Language and Gesture*, Cambridge 2000, pp. 284–311; David McNeill: *Hand and Mind*, Chicago, London 1992.

3 See McNeill: *Hand and Mind* (note 2), p. 15; Evelyn McClave (1994): Gestural beats: The Rhythm hypothesis. *Journal of Psycholinguistic Research* 23/1 (1994).

4 See e.g. Adam Kendon: Do gestures communicate? A review, in: *Research on Language and Social Interaction*, 27/3 (1994); but also Justine Cassell/David McNeill/Karl-Erik McCullough: Speech-Gesture Mismatches: Evidence for One Underlying Representation of Linguistic and Non-Linguistic Information, in: *Pragmatics and Cognition*, 7/1 (1999), pp. 1–33.

5 See Kendon: Do gestures communicate? (note 4).

6 De Ruiter: The production of gesture and speech (note 2).

7 Cassell/McNeill/McCullough: Speech-Gesture Mismatches (note 4).

8 Bernard Rimé/Loris Schiaratura: Gesture and speech, in: R. Feldman/Bernard Rimé (Eds.): *Fundamentals of nonverbal behavior*, Cambridge 1991, pp. 239–281; Robert M. Krauss/Palmer Morrel-Samuels/Christina Colasante: Do conversational hand gestures communicate?, in: *Journal of Personality and Social Psychology* 61/5 (1991), pp. 743–754; Robert M. Krauss/Yihsiu Chen/Rebecca

F. Gottesmann: Lexical gestures and lexical access: a process model, in: McNeill: *Language and Gesture* (note 2), pp. 261–283; Uri Hadar/Brian Butterworth: Iconic gesture, imagery and word retrieval in speech, in: *Semiotica* 115 (1997), pp. 147–172.

9 Krauss/Chen/Gottesmann: Lexical gestures and lexical access: a process model (note 8).

10 Ibid.

11 See Willem J. M. Levelt (ed.): *Lexical Access in Speech Production*, Cambridge, MA 1991, for a detailed discussion of theories of lexical retrieval.

12 McNeill: *Hand and Mind* (note 2); De Ruiter: *The production of gesture and speech* (note 2).

13 Palmer Morrel-Samuels/Robert M. Krauss: Word familiarity predicts temporal asynchrony of hand gestures and speech, in: *Journal of Experimental Psychology: Learning, Memory and Cognition*, 18/3 (1992), pp. 615–622.

14 See e.g. Jörg Jescheniak: *Word Frequency Effects in Speech Production*. Doctoral thesis, Catholic University of Nijmegen 1994.

15 Robert M. Krauss/Palmer Morrel-Samuels/Christina Colasante: Do conversational hand gestures communicate?, in: *Journal of Personality and Social Psychology*, 61/5 (1991), pp. 743–754.

16 Akiba A. Cohen: The communicative functions of hand illustrators, in: *Journal of Communication* 27 (1977), pp. 54–63.

17 Jean A. Graham/Michael Argyle: A cross-cultural study of the communication of extra-verbal meaning by gestures, in: *International journal of psychology* 10/1 (1975), pp. 57–67.

18 Jean A. Graham/Simon Heywood: The effects of elimination of hand gestures and of verbal codability on speech performance, in: *European Journal of social Psychology* 5/2 (1975), pp. 189–195.

19 Bernard Rimé/Loris Schiaratura/Anne Ghyselinckx: Effects of relative immobilization on the speaker's nonverbal behavior and on the dialogue imagery level, in: *Motivation and Emotion*, 8(4), 1984, pp. 311–325.

20 Graham/Argyle: A cross-cultural study (note 17); Graham/Heywood: The effects of elimination of hand gestures (note 18).

21 Franz B. Rauscher/Robert M. Krauss/Yihsiu Chen: Gesture, speech and lexical access: The role of lexical movements in speech production, in: *Psychological Science* 7 (1996), pp. 226–231.

22 Geoffrey Beattie/Jane Coughlan: An experimental investigation of the role of iconic gestures in lexical access using the tip-of-the-tongue phenomenon, in: *British Journal of Psychology* 90 (1999), pp. 35–56.

23 Alan S. Brown: A review of the tip-of-the-tongue experience, in: *Psychological Bulletin* 109 (1991), pp. 204–223 (here: p. 204).

24 Beattie/Coughlan: An experimental investigation of the role of iconic gestures (note 22), p. 49.