THEORETICAL NOTES

Gestures and Speech, Interactions and Separations:
A Reply to McNeill (1985)

P. Feyereisen
University of Louvain, Brussels, Belgium

Data from aphasiology and from developmental and experimental psychology are reviewed that give
evidence for partial separability of the gestural and the verbal systems. According to McNeill (1985),
who argued for a shared computational stage at the conceptual level, these cases of dissociation may
be accounted for only by assuming separation at the output level. This explanation seems incompati-
bile with other data provided by McNeill that indicate similarities in the surface characteristics of
gestures and speech. Thus it is concluded that gestures and speech could also interact on an interme-
diate level between conceptualization and the output processing or that the shared computational
stage must be specified by assuming separate subprocesses that may dissociate.

There are different ways to conceive the relations between gestures and speech. A traditional view holds that body move-
ments and language constitute parallel separate systems: Gestures are specialized in the expression of connotative meanings
and affective states, as are other nonverbal signals like eye movements and facial expressions (see Feyereisen & de Lannoy,
1985, for a review). According to this view, the variables controlling the gestural processes are independent of those involved
in speech production processes, and the two mechanisms are not connected. McNeill (1985) challenged this perspective by
arguing for a close association of the systems controlling verbal and gestural outputs: “Gestures share with speech a computa-
tional stage; they are, accordingly, parts of the same psychological structure” (p. 350). In support of this claim, McNeill noted
that gestures occur only during speech in synchronization with linguistic units, they serve similar pragmatic and semantic
functions, there is simultaneous dissolution of linguistic and gestural abilities in aphasic subjects, and gestures develop to-
gether with speech in children. The reason for such a strong connection might be situated in the nature of the thinking pro-
cess and its origin in the sensorimotor stage of intelligence ontoge-
nesis. “The basis for synchronization is not that gesture and sentence are translations of one another, but that they arise
from a common cognitive representation” (p. 353). (See Mc-
Neill, 1979, for a full development of the idea).

There is little dispute over the general claim that gesture produc-
tion depends to some extent on mechanisms that are also responsible for speech production. From Wundt (1900/1973)
to more recent reviews on gestural behavior (e.g., Feyereisen &
Seron, 1984; Kendon, 1983; Rimé, 1983), several hypotheses have been formulated about the interactions of the verbal and
gestural systems. Thus the problem is not just to demonstrate connections between these systems but to specify the way in
which they are connected. McNeill suggested that speech and gestures are different manifestations of a global synthetic represen-
tation called inner speech. Unlike models that assume, with differences in labeling the processing units and in charting the
flow of the information, a decomposition of the process of speech production in several subsystems each serving a specific
operation (see e.g., Bock, 1982; Dell, 1985; Garrett, 1982; Har-
ley, 1984), McNeill’s model relies on a single complex representation underlying the conceptual structure of the message, the
surface aspects of the utterance, and the functional and the morphological characteristics of the gestures. Gestures and speech
result from this common computational stage and separate into different outputs. The difficulty here is to define the processes
belonging to the common stage in which interactions occur and
to the later output processing in which dissociations may be
observed. Either the output processes refer only to the control
mechanisms of specific effectors (hands and articulatory sys-
tems), or the output processes encompass all the modality-spe-
cific operations, for example, in the verbal channel, giving a
phonological interpretation to the syntactico-lexical represent-
tation. To identify the computational stage shared by the ges-
tural and the verbal systems, two kinds of data may be used:
first, the description of the interactions between gestures and
speech, which allows the common computational stage to be
specified, and second, the analysis of the cases in which the two
systems separate or, in McNeill’s perspective, the cases in which
output processes are selectively involved.

Gestures and Speech: How Do They Interact?

Gestures accompanying speech could, according to McNeill
(1985), be “direct manifestations of the speaker’s ongoing

P. Feyereisen is a Research Associate of the National Fund for Sci-
cientific Research, Belgium.

Correspondence concerning this article should be addressed to P. Fey-
ereisen, University of Louvain, Unité NEXA, UCL 5545, Avenue Hippo-
crate, B-1200 Brussels, Belgium.
thinking process" (p. 367) and relate to the "initial phase of sentence generation" (p. 370). "At this level, two kinds of thinking are being coordinated: imagistic (thinking that is global and synthetic) and syntactic (thinking that is linear and segmented)" (p. 370).

This is not, however, the only way by which gestures and speech may interact. Gestures can also reflect difficulties experienced by the speaker in speech production and relate to the mechanisms controlling the verbalization processes (feed-forward and feedback loops). I suggest that an overload in working memory or a mismatch between the intended sentence and the utterance as planned may trigger gestures. Thus gestures would not be produced at any moment, but in reaction to specific obstacles encountered in the speech planning. This hypothesis is supported by the high incidence of hand movements during pauses in the execution (i.e., fluent) phases of the speech production process (Butterworth & Beattie, 1978), in bilinguals using the nondominant language (Marcos, 1979; Sainsbury & Wood, 1977), and in aphasic subjects (Feyereisen, 1983). That hand movements are more frequent at the end than in the beginning of a conversation (Feyereisen, 1982) also suggests feedback from the speech-planning system rather than a single connection at the conceptual level. Lastly, the speaker could sometimes rely on both the gestural and the verbal channels simultaneously to express related meanings, and in this way, too, gesture production would be influenced by monitoring processes in speech planning (Kendon, 1983).

Another possible interaction between speech and gesture is suggested by the notion of coordinative structures presented by Kelso, Tuller, and Harris (1983). They found finger movement amplitude to be related to stress in the utterance of lists of words, as though the motor programs controlling the manual and the vocal movements were exchanging information about the values of the parameters specifying details of execution. These observations were partially replicated by Smith, McFarland, and Weber (1986). It would be interesting to examine in more natural conditions whether some gestures (especially "beats") relate to the prosodic features (stress and melodic contour) or to the syllabic structure of the verbal utterances. The synchronization of manual and oral movements has been experimentally analyzed by Levelt, Richardson, and La Heij (1985). These researchers contrasted two hypotheses on the coordination of pointing with deictic utterances: full autonomy (relation during the planning phase only, but no feedback from one system to the other during the execution of the motor programs) and interdependence on the lowest level. The subjects were asked to point to lights turned on near or far from a center line while saying "this lamp" or "that lamp." As expected, the initiation time of the gesture, the apex time (time to reach the maximum extent), and the execution time (difference between apex and initiation times) were influenced by distance, the longer movements taking more time to be planned and executed. But the voice onset time also varied with the experimental conditions, thus revealing interaction between the two systems. Similarly, the voice onset time was influenced by the relative location of the light and the responding hand, the time being faster in the ipsilateral than in the contralateral condition. These variables had no effect in another task in which the subjects responded only verbally. However, an unexpected application of a load on the arm after the initiation of the pointing movement delayed the apex time (as one may expect) but did not influence the voice onset time. Thus, as predicted by the hypothesis of autonomy, interaction between the gesture system and the speech system was absent during the execution phase and restricted to the phase just prior to beginning the movement. Gestures and speech may be characterized as ballistic movements rather than movements subject to continuous control by feedback. The problem remains to specify the nature of the interaction during the planning phase (competition for common resources, coordinative structure) and to establish whether it occurs at an early stage (conceptual level: spatial encoding of the stimulus light, selection of the response) or in the late processing (specification of the parameters of the movements). Another illustration of the conception of coordinative structures is also provided by the description of head movement in speech production (Hadar, Steiner, Grant, & Rose, 1983a, 1983b, 1984). Hadar et al. suggested a possible role of these movements in the phasing of articulatory movements or a coordination of the control mechanisms for the head musculature and the speech apparatus. Parallels with hand movements remain to be studied.

Thus interactions between gestural and verbal systems could occur at different levels of processing: common underlying representations of action at the conceptual level, matching of the lexical selection output with the source message, coordination in the motor planning for speech and hand movements, and so on. By assuming a shared computation of gestures and speech, McNeill was probably referring to all these operations; the different kinds of possible interactions between the gestural and the verbal systems would be, from his point of view, merely different facets of the thinking process generating visible or audible outputs. An unfortunate implication of this conception is that a heavy cognitive load is imposed on the computational stage of programming manual and vocal movements. Moreover, little opportunity is given to observe dissociations between gestures and speech unless by addressing to the output processes the variables that influence one system and not the other.

When Gestures and Speech Separate

Separability is classically evidenced by pathology when one system is disrupted and the other is spared, by the study of development when "décâlages" are observed in the emergence of different functions, and by experimental manipulation of variables affecting one system and not the other. McNeill (1985) stressed the idea that "speech and gesture respond to the same factor at the same times" (p. 350). Arguments supporting this view derive, for example, from the behavior of aphasic subjects and of children. In these two areas, McNeill referred to Bates, Bretherton, Shore, and McNew's (1983) review, which attempts to demonstrate similarity in the use of vocal and gestural symbols and concludes with the association of linguistic and gestural behavior in development and in pathology.

The data on aphasia, however, are disputed (Feyereisen & Seron, 1982, 1984). There is, first, a discussion about the meaning to be given to the association of verbal and gestural deficits in left-hemisphere-damaged patients: Is it a spurious correla-
tion attributable to the simultaneous impairment of distinct control centers, or does the association indicate the existence of common, underlying mechanisms? If the latter, the question about the nature of the impaired mechanism arises: Are gesture and speech disturbed because of incomplete specification of the input at the conceptual level or because of disrupted control of the motor programs (Kimura, 1976)? The relevant data are derived from the analysis of ideomotor apraxia and manual activity during speaking.

Ideomotor apraxia is the inability of some brain-damaged patients to perform symbolic gestures on request (verbal instruction or elicited imitation), such as conventional manual signs or pretended object use. This gestural impairment is often observed in association with aphasia, and correlations between severity of praxic and linguistic disorders provide the main empirical support for the hypothesis of a central symbolic deficit (e.g., Duffy & Duffy, 1981). However, high correlations in pathological populations are not enough to demonstrate the existence of an in Central Mechanism controlling both verbal and gestural production. A single case of dissociation would be sufficient proof of the separation of the underlying mechanisms. Indeed, aphasics can often pantomime the use of objects they are unable to name (Davis, Artes, & Hoops, 1979). Aphasia without apraxia is common, and Kertesz, Ferro, and Shewan (1984) have described the anatomical basis for such dissociation. The phenomenon is also reported among deaf signing patients, who produce language in the gestural modality (Poizner, Bellugi, & Iragui, 1984). To some extent, producing nonlinguistic movements on request is a simpler task than producing linguistic signs, be they oral or manual. A related observation is the advantage of sign languages over oral languages in the acquisition of communicative skills by mentally handicapped children (Abrahamsen, Cavallo, & McCluer, 1985). The inverse dissociation of apraxia without aphasia is much less frequent when verbal behavior is carefully assessed, but it is sometimes observed: five cases have been reported in the series of de Ajuriaguerra, Hécaen, and Anjerluegos (1960), and two cases in the series of de Renzi, Motti, and Nichelli (1980). A single case was described by Heilman, Gonyea, and Geschwind (1974). Separation of the underlying mechanisms has also been observed in left-handed people recovering from aphasia, but not from apraxia (Selnes, Rubens, Risse, & Levy, 1982). Other arguments against the hypothesis of a unitary disturbance of language and gesture are provided by evidence that the modality through which gestures are elicited—verbal, visual, or tactile information—influences the performance of some patients (De Renzi, Faglioni, & Sorgato, 1982) and also by the loose association between gestural impairment and subtypes of aphasia (Kertesz & Hooper, 1982; Lehmkohl, Poeck, & Willmes, 1983). Even if the kind of gestures disturbed in ideomotor apraxia differs in many respects from the spontaneous gestures analyzed by McNeill, it is unclear what these differences mean if the production of both gestures and speech is considered to derive from a common conceptual structure. Nevertheless, it could be that conventional or symbolic gestures dissociate from language in pathology, whereas spontaneous gestures accompanying speech do not.

The available data do not favor such a view. First, in the comparison of aphasics with normal subjects, spontaneous manual activity seems less disrupted than language, both quantitatively (Feyereisen, 1983) and qualitatively (Pedelt, cited by McNeill, 1985). McNeill concluded that “the occurrence of appropriate gestures implies that the paraphasias are not due to conceptual confusions, but arise from disturbances after the speaker has developed an appropriate cognitive representation” (p. 382). Butterworth, Swallow, and Grimston (1981) drew a similar conclusion. This raises the problem of the nature of the connection between gestures and language, which will be examined shortly. In the comparison of subtypes of aphasia, the major difference is not the number of gestures produced (Ciccone, Wapner, Foldi, Zurif, & Gardner, 1979; Feyereisen, 1983), but the referential value of the hand movements: Broca’s aphasics produce mainly referential gestures, called iconix by McNeill, whereas about 50% of the movements of Wernicke’s aphasics are considered beats (i.e., noninformative). The problem, however, is that the referential value of gestures cannot be assessed unambiguously on the basis of the formal characteristics of the movement. According to McNeill, 1985, these gestures “are not interpretable in the absence of speech” [p. 351]. “An iconic gesture is one that in form and manner of execution exhibits a meaning relevant to the simultaneously expressed linguistic meaning” [p. 354]. “Beats have no propositional content of their own... They are particularly appropriate for emphasizing discourse-oriented functions” [p. 359]. Recent empirical data also illustrate the influence of simultaneously verbally expressed contents on the classification of gestures (Feyereisen, Van de Wiele, & Dubois, 1987). Thus qualitative differences in the gestural production of Broca’s and Wernicke’s aphasics would reflect differences in verbal production, not because both depend on a common computational stage but because gesture description is influenced by the characteristics of concomitant speech.

Data from language acquisition also show both interactions and dissociations of the verbal and gestural systems. On the one hand, Bates et al. (1983) described analogous events in the second year of development of language and gestures. “In brief, vocal and manual names emerged around the same time and correlated across the sample in frequency of use, rate of acquisition, and numbers of different schemes” (p. 71). But, on the other hand, there are circumstances in which gestures and speech dissociate. First, gestures are strongly associated with object manipulation; their ties to objects hinder the decontextualization process and representational functioning. By playing with actual objects (realistic condition) or with toys schematically representing these objects (abstract condition), 13-month-old children produce gestural schemes more often and vocal schemes less often than do 20-month-old children. Thus the threshold for eliciting gestures may be lower than for words, which makes the gestures the preferred modality in younger children (Breherton et al., 1981). At both ages, more verbal and gestural productions are observed in the realistic condition, but in the gestural modality, the difference is greater at 13 than at 20 months, leading to a significant Age × Condition interaction. A similar interaction is not observed in the vocal modality. Thus the physical appearance of objects would have more influence on gestural than on vocal production. Rather than the unitary view advocated by McNeill, these observations favor domain-specific or local homology models accounting for correlations restricted to some tasks or to some stages of development (Bates, Breherton, Snyder, Shore, & Volterra, 1980). Fur-
Furthermore, the picture of the relation between gesture and speech in development could differ according to the kind of gestures considered. The gestural symbols described in early childhood may have a different basis from spontaneous gestures accompanying speech. If symbolic play precedes the occurrence of the corresponding word, iconic gesture follows it, which again shows the separability of the two systems. As far as pointing movements are concerned, there seems to be a relation not to acquisition of linguistic devices (Dobrich & Scarborough, 1984) but, rather, to less specialized cognitive abilities, such as dividing attention between people and objects (Leung & Rheingold, 1981; Masur, 1983).

Finally, there is some evidence for an experimental dissociation of gesture and language. Studies examining the influence of independent variables on both speech and gestures are not numerous, nor are studies devoted to the experimental analysis of the speech production processes. However, manipulations intended to influence speech output might have little impact on manual activity. For example, giving directions to proximate or distant places and thus producing utterances of probably different length does not change the use of gestures (Cohen & Harrison, 1973). Conversely, restriction of hand movements during speaking seems to have a limited influence on the linguistic output: It affects imagery index but not lexical variety (Rimé, Schiaratura, Hupet, & Ghyselenlincx, 1984), and it influences use of deictic expressions or of words describing spatial relations, but not the number of words or descriptive nouns (Graham & Heywood, 1975). Thus gestural or verbal behavior can be affected without parallel changes in the other modality, and some variables can influence one aspect of the production and not the other.

Conclusions

The problem now is to account for the data demonstrating separability of the gestural and verbal system and to explain, at the same time, the connection or connections between the verbal and the gestural systems. The solution proposed by McNeill is that there is a unitary computational level the results of which are processed in two separate output systems. This suggestion implies that dissociations can be explained by the characteristics of the output systems. Appropriate gesturing in aphasia, according to McNeill, argues for intact cognitive representation, flawless transformation into the gestural channel, and defective transformation into the vocal channel. Similarly, anteriority of gestures over speech in language development would indicate that manual activity is either more closely connected to the thinking process than speech or is a simpler, perhaps more direct, representation of the conceptual level. Thus gestures would truly reflect the cognitive state of the subject, whereas corresponding vocal utterances would be more sensitive to brain damage or could be delayed in development. In this line of thought, the properties of verbal utterances that distinguish aphasic from normal speech, children of different developmental stages, or experimental conditions in which gesture and speech dissociate all depend on the characteristics of the speech output system. But the similarities in the characteristics of gestures and speech also argue for a common processing in a shared computational stage. McNeill described the successive emergence in early childhood of three kinds of gestures—deictic, iconic, and beats—paralleling the appearance of the first spoken symbols, the progressive decontextualization of the expressed meanings, and the attainment of the text-coding stage. The differences between the three stages of development must therefore involve the shared conceptual level from which both gestures and speech originate. Similarly, gestural behavior of aphasic subjects would also manifest the characteristics of the different types of pathology: Broca's aphasias produce numerous iconic gestures, and their speech is "telegraphic" because it is informative but lacks grammatical words. Conversely, Wernicke's aphasias mainly produce beats as movements, and their speech, although syntactically connected, is said to be empty because it lacks content words. If the difference between these two aphasias syndromes is manifested in gestures as well as in speech, it must be located in a computational stage common to both systems (i.e., before the separation into distinct output processors). Thus there is a contradiction between two kinds of evidence: dissociation of the two control systems and similarities in the characteristics of speech and gestures. Is the language disorder in aphasia just a disruption of the verbal output system? If so, how can one explain similarities and the absence of compensation via the gestural output system (making Wernicke's aphasias produce referential iconic gestures and Broca's aphasias relational beat gestures)? Or is the conceptual level where gestures and speech are conceived itself disrupted? But, if so, how can one explain why gestural behavior seems less disrupted than verbal behavior? More generally, how can one reconcile separation and connection of the gestural and verbal systems?

The problem arises from a lack of specification of the cognitive representation underlying speech and gesture production, apart from a general, multifaceted process called thinking and identified with inner speech. An alternative to this view offers a better account of the available data by suggesting the decomposability of the mechanism underlying speech and gestures. Gestural and verbal production would result from separate but interacting processes. Within each system, several distinct operations would have to be distinguished. For example, in the verbal system, the process of mapping semantic representations into a syntactic frame might be dissociated from the process of accessing phonological representations from the lexicon. Accordingly, audible speech output would result from several parallel or successive recordings of the conceptual representation of the message. Similarly, gesture appearance would conform to motor parameters simultaneously or sequentially computed at several different levels. Hence multiple connections may exist between the two systems; synchronization of speech and gestures and similarities in function or in development may be due to associations at different stages of the production. Different kinds of gestures, or more precisely, different parameters of a given movement, may also relate to different processes in the coordinated systems. According to this alternative conception, cases of dissociation may be explained by assuming selective influence of some variables at a stage following the conceptualization phase. Another fortunate implication of this proposition is that it suggests interactions between gestures and speech at lower levels of processing than in the initial level of message
conceptualization and thus lightens the burden on this computational stage in programming vocal and manual movements.

The provocative question posed by McNeill (1985) has drawn attention to the puzzling behavior of people who gesture while they speak. It also suggests that studies in this domain could shed new light on the nature of the mental representations underlying speech production. Thus McNeill’s article should elicit more research on this topic. Unfortunately, the solution McNeill proposed leads one to believe that the problem is now solved (when, in fact, it is just being raised) and that we know more than we actually do. A more heuristic procedure would seem to confront alternative views on the interactions between gestures and speech: Either the gestural and the verbal systems are only connected on the conceptual level, or information can be exchanged during further elaboration of the message on the motor-planning level. Thus researchers must determine which computational stage is shared by the two systems and when they begin to diverge.

References


Received May 16, 1986

Revision received October 10, 1986

Accepted October 10, 1986

---


The Publications and Communications Board of the American Psychological Association announces the appointment of Neal Schmitt, Michigan State University, as editor of the *Journal of Applied Psychology* for a 6-year term beginning in 1989. As of January 1, 1988, manuscripts should be directed to

Neal Schmitt  
Department of Psychology  
Psychology Research Building  
Michigan State University  
East Lansing, Michigan 48824

Manuscript submission patterns for the *Journal of Applied Psychology* make the precise date of completion of the 1988 volume uncertain. The current editor, Robert Guion, will review and consider manuscripts until December 31, 1987. Should the 1988 volume be completed before that date, manuscripts will be redirected to Schmitt for consideration in the 1989 volume.