

## **Kinesic Components of Multimodal Utterances**

ADAM KENDON

*University of Pennsylvania*

### **0. Introduction**

I begin with a commonplace observation: when people speak they often (although not always) engage in hand and arm movements which, typically, are seen as “part of talking” - these hand movements are often deemed meaningful in some way and are commonly said, for example, to add emphasis to what the speaker is saying or to add to the description of something being talked about. These are normally recognized as movements that are intended to convey some kind of utterance related meaning, and in the ordinary course of our dealings with one another they rarely, if ever, are seen as strange or unintelligible.

Such movements, commonly referred to as “gestures” or “gesticulation”, in the last twenty years or so have become the focus of much scientific analysis. This work, especially that done on so-called co-speech gestures, has been driven in large part (so it seems to me) by a great *puzzlement* as to why people engage in them. As I have just suggested, however, this is not a puzzlement that ordinary people in their everyday lives suffer from (‘gestures’ are quite transparent, I think, for most people). It is a puzzlement that is suffered from by modern psychologists, psycholinguists, and others. This is in part because we have been, at least until recently - and perhaps still are - under the very strong impression that anything we need to say to another, at whatever level of abstraction and complexity it might be, ought all be done in *words* - and we really don’t need anything else. If this is so, why use gestures at all? Perhaps, as is sometimes suggested, gestures are only used when ability with words is lacking. For this reason, a negative value has often been attached to them. Sometimes there is even hostility towards gestures - using them indicates a lack of proper self-control as well as a lack of proper mastery of language, their use shows “lack of breeding” and therefore they should be suppressed.

The idea that verbal language is sufficient for all communicative purposes (with the underlying attitude that if it isn’t, it ought to be) that remains dominant even today, is quite old, but was reinforced especially, I think, by the rise of mechanized printing technology in the nineteenth century, which finally estab-

lished the *written word*, language in its *written form*, as the ideal to which we, as speakers should all aspire and, indeed, as the most perfect and complete expression in language possible. Some have argued that the course of modern linguistics has been deeply shaped by this (see Harris 1987, Linell 2005). Linguistics, it is suggested, is largely the study of language in its written form, and where this is not so, it nevertheless remains that linguistics is mainly concerned with what, from speech, *can* be written down with common writing techniques. Of course, many aspects of what we do with our voices when speaking - including intonation, for instance - are being incorporated into linguistic science, but this incorporation began only relatively recently and it is far from complete. As for kinesic aspects, including so-called “co-speech gesture”, these are only grudgingly admitted, if at all. Co-speech gesture, in particular, though now it attracts considerable attention, does so not so much for its own sake but for what light it might throw on the *speaking* process. Thus we wonder: do these manual actions help us retrieve the words we need to speak with? (e.g. Krauss et al. 2000) or do they help us organize our thoughts better so that we can speak them? (Kita 2000, Goldin-Meadow 2003) or is it that they make manifest the imagistic aspects of thinking and so, by studying them, we gain a special ‘window’ on to what is going on in the mind of the speaker as he speaks? They can be useful, it is argued, as indicators of what is involved in what has been called *thinking for speaking* (Slobin 1987; McNeill 2005).

Questions of this sort, I suggest, would only have arisen against the background of a view of language and speech that sees it as strictly *verbal*, and sees it, in fact, as what can be abstracted from the flow action that counts as utterance largely, perhaps wholly, in terms of what we can write down. And this view of language, so I think, finally came to dominate after written language came to dominate the culture, as it did, finally, in Europe, at least, in the nineteenth century. If you go back to the eighteenth century, and earlier, the role of gesture in discourse seems much more accepted, indeed it was cultivated as a part of an Art of Speaking (see, for example, Austin 1806; see also Barnett 1987). But as European culture finally lost its oral character and became dominated by language in its written form, those who thought about how language operates, the process of communication, and so forth, thought only about how this worked through words. It took the revolution in recording technologies to remind us (rather to many people’s surprise, in fact) that in communicating with one another people also employed their bodies in a variety of ways. It was as a consequence of this that the idea of “nonverbal communication” became current (Kendon 2004, ch. 5). As may be noted, this term pre-supposes that there is the communicating that we do with *words*, on the one hand, and then there is the other kind. The very term encouraged an approach in which it was to be accorded separate treatment. Only very lately, after our culture has been saturated for many decades with cinematic, video and television imagery, is the idea at last becoming widespread in learned circles that communication that involves language actually also involves more than the words spoken. Nowadays, one does not hear the term “nonverbal com-

munication” quite as often. Nowadays it is recognized that communication is “multimodal”. Steps are beginning toward a more integrated view.

As I have already suggested, as a consequence of this history, the twenty years or more of growth in the study of co-speech gestures has mainly been about what they might reveal about internal psychological processes. As a result, we largely lack good detailed descriptions of how, in the production of utterances, these movements are used and how they may contribute to the utterances of which they are a part. We have few accounts of the forms of gestural expressions that speakers commonly use. We lack accounts of how the hands are used as descriptive devices or how they express the pragmatic aspects of a participant’s engagement in talk (as they commonly do), what the circumstances are in which gestures are used and what the circumstances are when they are not used, whether and in what way there are changes in how gestures are used as the speaking register changes, the extent to which there are cultural and social class differences in the use of gestures - and so on. We remain with very little systematic understanding of most of this.

### **1. Different Uses of Manual Actions by Speakers**

In this talk I will present three very short extracts from different occasions of rather commonplace interaction situations in which the speakers are using gestures, and with these examples I intend to illustrate some of the principal ways in which, as far as I can see, such manual actions enter into and operate as a part of the utterance construction process. I shall do this entirely from a production point of view. In the examples I shall present it will, I think, be obvious, that these movements are completely integrated into the *ensemble* of actions which constitute the utterances in question. They are part and parcel of the whole way in which the speaker goes about producing the complex of acts intended *for* another or others as an utterance, the complex ensemble of actions by which he *says* something to another or others.

Whether, and to what extent, other persons actually process these movements and the extent to which this contributes to other persons’ perception or comprehension of what the speaker is doing is not a question I shall enter into here. There are good reasons to suppose that the kinesic components of utterances do have consequences for how an utterance is perceived or responded to, but I won’t review this here (see, for example, Beattie and Shovelton 1999, 2001, 2002). My interest, rather, is to try to give some concrete illustrations of the way in which people engaged in utterance production deploy their speech and kinesic actions in relation to one another and how, in doing so, kinesis and speech are employed together in a kind of constructional process. As my three examples will show, there are several different ways in which this happens - these “semantically significant movements” or gestures can contribute in *different* ways to the utterance of which they are a part.

The three examples I shall present are very short and utterly commonplace. Nevertheless, as you will see, they are ‘dense’ in gesture use and show consider-

able richness. All three of them are drawn from recordings made in Italy - two in or near the city of Naples, one in an academic context at a University in Calabria. They have been selected to illustrate three of the main ways in which manual actions appear to be used by speakers.

First of all, manual actions recognized as gestures, may be employed in such a way as to provide dimensional and dynamic movement information about the properties of objects or actions the speaker is talking about. In this usage, these movements often have a kind of adjectival or adverbial function, giving an idea of the relative dimensions, shape or spatial positioning of nominated objects, or referring to the manner of some action mentioned by a verb. When used in this manner, such movements can contribute to the propositional content of what the speaker is saying. They play a role that could have been played by words, had the speaker cared to use them.

Another common use of manual actions while speaking is to use them as a way of suggesting the presence of concrete objects which are offered as representations of the entities (these may be abstract or concrete) a speaker is talking about. Such entities, once created through manual action, can then be moved about or placed in relation to one another and so serve as visual diagrams or movement patterns that display in visible form aspects of the relationships obtaining between these entities also being described in the associated spoken discourse. In this sort of usage, the hand movements may not add anything to the propositional content of what is being said, but they may represent this content in another way. In these cases the gestures work much as diagrams or drawings do that are used to exhibit visually things being explained in a written text.<sup>1</sup> Both of these kinds of uses may be called *representation* uses of manual actions.

It also is common for speakers to engage in manual movements that can be recognized as schematic forms of actions such as offering, grasping, withdrawing the hands, holding small objects in a precision grip, sweeping things aside or brushing things away, and the like, which are so placed in relation to the speaker's verbal discourse that they are understood as expressions of the forms of *pragmatic action* that a speaker undertakes in speaking. Here, it seems, by these manual actions the speaker acts out or performs, albeit in a highly schematized manner, the illocutionary actions performed in the talk. As speech act theory proposes, any act of speaking also counts as a mode of action, a mode of making a move of some sort in respect to those for whom the act of speaking is intended. Thus in saying something one may *assert, request, beg, deny, withdraw, hold up, stop, offer, present, indicate* and a host of other actions. One often finds that the manual actions associated with speaking are forms of action that express these kinds of

---

<sup>1</sup> Such gestures sometimes are called "metaphoric gestures". The gestures, of course, are not metaphoric although the treatment in the speaker's discourse of abstract ideas as concrete objects or processes which the gestures illustrate is. For discussion and examples of this sort of gesture use see, especially, Cienki and Müller (2008).

pragmatic meanings rather than expressing something about the propositional content of the discourse.

In fact, this is an extremely important aspect of gesture use which, quite surprisingly, has received very little attention in contemporary gesture studies. This sort of gesture use was well known to students of rhetoric, and in one of the more famous passages from Quintilian's treatise on Roman rhetoric from the First Century AD, gestures of this type are given detailed treatment (Butler 1922: XI, line 85ff). For detailed modern discussions, see Kendon 2004 (ch. 12 and 13) and Streeck (2009).

Two other important uses of gesture must be mentioned, although these will not be illustrated here. There is the use of gesture in *pointing* and the use of gesture in *interaction regulation*. So far as *pointing* is concerned, people use their hands to indicate the objects that are the topics of their discourse and, as is very well known, sometimes a piece of spoken discourse is incomprehensible without the kinesic deixis that is conjoined with it. Pointing is itself a complex phenomenon and can be done in many different ways, and the way it is done - for example, what sort of a hand shape is employed, or whether a pointer uses head or lips or hand or foot, and so on, - may have semantic consequences of various kinds (See Kendon 2004, ch. 11). With regard to *interaction regulation*, there are many ways in which manual actions are employed for this. For example, manual bodily action is involved in the offering and exchange of *salutations*, one sees people directing people to engage in actions with their hands - such as 'stop' 'approach' 'go back' 'turn round' and so forth. Or there are ways in which people employ their hands to claim or to request a turn at talk in conversation.

This division into four uses of gesture is not, I should emphasize, intended as a *typology* of gesture. It is a useful guide, however, at a very broad level, to the main ways in which gestures are used. It is important to see that any given instance of a gestural act may very well be participating in more than one of these uses at once - although it seems that gestural actions can be more or less *specialized* for these functions. Especially when gestures are used for pointing, interactional regulation and pragmatic purposes, we find there a tendency for forms to be conventionalized. Much gesturing used for representation may also follow conventional forms, but if you think about the vocabularies of so-called "emblems" that have been published, an analysis of what they are reputed to mean shows a preponderance of interaction regulation and pragmatic functions, rather than representation functions (see Kendon 1981, 1984; Payrató 1993).

## 2. Three Examples

I turn now to my three examples. In the first example we shall see how a speaker, in describing something, makes use of hand movements as part of the description, so here the hand movements enter directly into the utterance as a description and they add to the propositional content of what the speaker is saying. In the second example, we shall see a speaker who uses his hands to create a visual representation of the entities he is talking about and to show how these entities are related.

His use of hand movements here is not to add propositional information to what he is saying. Rather it is to present the contents of his talk in another mode. In the third example we will see illustrated how a speaker employs his hands to act out the pragmatic moves of his talk. In actions of this sort he makes manifest his pragmatic moves in kinesic form.

### **2.1. Example 1**

The first example is taken from a video-recording made some years ago in a small town near Naples where a group of young people were talking about various things, including telling about the pets they owned. Here a girl has said that she has an aquarium. One of the adults present asks her to describe this, and the extract I discuss is the first part of her description.

In response to the request to describe the aquarium, the speaker begins a description of her aquarium with three tone units. She says:

- (1) *Eii ci sono le pietre piccole/ci sono due pietre grandi/e ci sono l'erba là*  
Eii there are the little stones/there are two big stones/and there is the grass

Each tone unit mentions a component of the aquarium – the gravel (“*le pietre piccole*”), the rocks (“*due pietre grandi*”) and the vegetation, which she here calls “the grass” (“*l'erba*” although she means to say “*alghe*” – she corrects herself in a phrase immediately following which is not shown here). Associated with each one of these three phrases, distinctive manual actions can be seen. These we can analyze into what I call “gesture phrases” and in these terms there are two gesture phrases associated with the first unit, one with the second, and two with the third.

First, as she refers to the “*pietre piccole*” (the gravel in her aquarium), in her first gesture phrase she lifts both her hands up, held open palms down, moves them away from one another in a horizontal fashion, following this immediately by a second gesture phrase, accomplished by her right hand only, and done just as she says “*piccole*” in which the thumb and index finger are extended and held so that their tips are opposed. As I will mention again later, this gesture is often used as an expression that is the equivalent of the word “piccolo – small”. Almost everybody sees the double hand movement of the first gesture phrase as a depiction of the spatial disposition of the “*pietre piccole*” – that is to say, with this horizontal double hand movement she is taken to suggest a horizontal distribution of the gravel she is referring to. I will come back to comment on the gesture performed with “*piccole*” in a moment.

Second, as she says “*poi ci sono due pietre grandi*” she lifts both her hands up together, moves them outwards into the space in front of her at a level somewhat higher than her previous gesture. The hands are held with the palms facing outwards, the thumb and forefinger of each hand are extended, opposed but flexed, the other digits flexed toward the palm. This action is typically seen by most people as an action that serves to place or position two objects in front of

her, and it is taken to be the two “*pietre grandi*” that this hand action refers to, the action as it were displaying the placement of these two large stones.

Third, as she says “*e sono l'erba là*” she brings her two hands in front of her, oriented with palms down, thumb and forefinger brought in close opposition to one another (though in a relaxed manner), and then rapidly lifts first right hand then left hand in alternating actions. Just as she says “*là*” at the end of the phrase, her two hands turn outward and are held palm out with thumb and forefinger extended. The alternating upward lifting hand action of preceding gesture is typically seen by observers as a sort of “pulling upwards” action and is taken to mean that something is shown standing vertically – not a rigid thing, however; and the alternating actions of the hands suggest that there are several of them. Combined with “*l'erba*” it is typically understood as an action that refers to the vertically upstanding vegetation inside the aquarium.

In short, it is as if, with these actions, the speaker is sketching in certain visual features of the objects she is referring to. If we accept this, we can see that in such sketching she provides certain additional information that is not to be derived from her words. We get the idea, from the combination of the horizontal movement of open palm-down hands with the first phrase, that the little rocks or gravel she is referring to are distributed horizontally; from the combination of hands held forward of the third gesture with the phrase referring to the “*due pietre grandi*” we are given the idea of two objects – here the “*pietre*” – placed side-by-side one another above the horizontal gravel; and from the combination of the right-left alternation of upward movements of the two hands held in a loose thumb-finger precision grip configuration, together with a reference to “*l'erba*”, we are given the idea of the upward standing filaments of “grass” or “algae”.

I think seeing the gestures in this way and understanding the contribution they make to the overall meaning depends a good deal upon knowing that the speaker is talking about an aquarium and upon our general knowledge of what small aquariums are like. Nevertheless, I think it is clear that with these gestures, placed as they are in relation to her words, we learn not only *what* components of the aquarium there are but we learn something of how they are disposed in space in relation to one another and we are referred to what an aquarium might look like. Words and manual actions combined, thus, if the manual actions are understood in the way we have suggested, provide a *richer* description of the aquarium than just her words alone would do. This kind of use of manual action when speaking in discourses in which objects are being described is entirely typical. There is no need for me to multiply examples of this sort of thing here. (See Kendon 2004, chapters 8-10 for many similar examples. See also McNeill 1992 and Streeck 2009.)

Now, if we accept the interpretations I have given of these hand movements - that is, that they are semantically intended movements, the first one showing ‘horizontalness’, the second expressing the concept of ‘little’, the third showing two largish objects placed in space and the fourth showing a plurality of flimsy upstanding things - we can see that, when taken in conjunction with the words

uttered, they provide information about the spatial disposition, arrangement and visual appearance of the objects in the aquarium being described, well beyond anything that could be gathered from the words alone. They are a part of the speaker's description, that is to say. The features suggested by the hand movements could equally well have been suggested by some additional words. She could have said, for example, "*Eii ci sono le pietre piccole disposte orizzontalmente sul fondo dell' acquario, poi sopra ci sono due pietre grandi, una accanto all'altra, e ci sono l'erbe là, che crescono in modo verticale* - Eii there are the little stones arranged horizontally at the bottom of the aquarium, then there are two big stones above, side by side, and there is grass growing upwards vertically." But she didn't. She used hand movements of a certain sort to refer to some of these aspects, instead.

The speaker, in the turn space she is given, is to convey some idea of how her aquarium is organised, and she does so using words to nominate certain objects and hand movements to say something additional about how these objects are disposed in space and what they look like. How these hand movements serve to do this is by using visuo-kinesic schematic devices which serve to evoke or enable concepts like 'horizontalness', 'solid objects placed in space', 'flimsy strips rising upwards'. Thus, her two-handed horizontal gesture is a conventionalised schematic that stands in relation to the concept 'horizontal' in much the same way that the word itself does. The hand poses in the gesture done in relation to "*due pietre*" is a widely used device to suggest 'solid object' and the dynamic character of the gesture here is a widely used *practice* in kinesic expression to convey the idea of something firm in its position in space. In other words, these hand movements are *kinesic schematics*, constructed according to certain widely shared kinesic expression practices which refer us to general conceptual categories about shape, spatial position, and size (cf. Streeck 2008). As such, they can be employed in the task of utterance construction, playing a role very similar to that of verbal lexical items. Sometimes these kinesic schematic actions can be fully conventionalized and semantically equivalent to spoken words. We see this in the movements performed as she says "*piccole*" or "*là*". These are not only timed coincidentally with the pronunciation of these words, they are commonly glossed with these words. The gesture she does when she says "*piccole*" is a conventionalised form, widely used by Neapolitans in certain contexts when they refer to something that is "small" or "little" or "a little bit" (see de Jorio 2000:329-330). When she says "*là*", as already noted, she extends her index fingers forward, as if pointing, an action semantically equivalent to saying "there".

A further point may now be noted. The movements associated with the words "*piccole*" and "*là*", as we have just noted, seem to be kinesic equivalents of just those specific words. Here these movements really do seem to have a status just like that of individual lexical items. On the other hand, the horizontal hand movement done with "*ci sono le pietre* - there are the stones", the extended-arm-hand in 'grasp' pose done with "*ci sono due pietre grandi* - there are two big stones" and the alternating upward hand movements done at the same time as "*ci*

*sono l'erba* - there is the grass” extend over several words and, furthermore, they do not fit semantically with any words in the verbal component. Here their meaning seems to add to the meanings expressed in the words with which they are associated. Yet, as we have suggested, these movements are kinesic schematics and refer to conceptual categories, very much like spoken words do. We can see them, then, as entering into the utterance construction as if they might be lexical items (working as adjectives or adverbs), yet, because they are *kinesic* forms they can be done (as they are here) simultaneously with speaking and can, accordingly, add further semantic expression *simultaneously* with the words. These movements don't have to *alternate* with words, so they don't enter into the organisation of the talk sequentially, but this does not seem to be a good reason to exclude them as lexical elements in the utterance construction. Rather, we may say that the speaker's sentences are being constructed in more than one dimension at the same time. The speaker here shows us how it is possible to engage in what have been termed “simultaneous” linguistic constructions.

The notion of simultaneous linguistic constructions, which might require a representation of syntactic relations in three, rather than two dimensions, has recently been discussed by Bencie Woll in an essay dealing with a debate (Woll 2007), that began in the eighteenth century, concerning the apparent fact that for speakers the meaning of what they want to say is present “all at once” and that a problem they face, if they are to express this in spoken language, is to organize the material so that it can be presented in the linear manner that speech requires. It is the function of syntax to serve this linearization of expression. However, it has been suggested that this is a consequence of the kind of expressive vehicle that speech makes use of. It uses a set of articulators which serve to produce successions of sounds and it is this that forces speech to be constrained to the temporal linear dimension. In sign languages, on the other hand, the hands and the head, the face, the mouth, and also the whole body, can all serve as articulators and this means that there can be a differentiation of expression in space as well as in time. And, as students of sign languages are coming to see, *simultaneity* in syntactic construction is not only a possibility but is actually observed. It has been noted as a feature of signed languages since the work of Stokoe (1960), and has lately become the focus of a number of investigations (Vermeerbergen et al. 2007).

Woll describes the contrast that is often made between signed languages and spoken languages in this regard: “with just one set of articulators, spoken languages have linearity; with multiple articulators (two hands, face, mouth, etc.) signed languages have simultaneity” (Woll 2007:340). However, it should be pointed out that users of spoken languages *also* have multiple articulators. As we see in the case of the example we have just examined here, making use of the hands as our speaker did in this case, she was able to overcome the limitations of the spoken modality and was able to produce a construction every bit as ‘simultaneous’ as any we might find in signed language discourse.

I am suggesting, in other words, that at least in the case of this kind of gesture-use in relation to speech, we can entertain the possibility that a speaker may

employ units of expression that have semiotic functions quite similar to spoken lexical expressions which, nevertheless, being in the kinesic modality can be employed simultaneously and in parallel with words. The speaker thus has available a solution to the constraints that a strictly spoken verbal modality imposes.

## **2.2. Example 2**

I turn now to Example 2, which shows a rather different way in which manual expressions can be employed by a speaker. This example is taken from a video recording of a seminar in which the speaker was discussing aspects of Down's syndrome. In this example we see several further features of the gesture/speech relation and we see another way in which a speaker can use manual expressions in relation to his verbal discourse.

In the extract considered here, the speaker mentions two concepts: cognition and language. He refers to the observations that some have made on Down's syndrome children that, because these children may show a good deal of linguistic sophistication, yet are apparently retarded in respect to other kinds of cognitive abilities, this is a natural demonstration that "cognition" and "language" are separate components or modules, that linguistic ability is independent of general cognitive abilities. He says:

- (2) *questa è la dimostrazione naturale/che la cognizione e il linguaggio/  
sono aspetti dissociabili tra loro/no?/perché non c'è una correlazione*  
this is the natural demonstration/that cognition and language/  
are aspects dissociated from one another/no?/ because there is no correlation

For the purposes of our immediate interest, we begin with his second phrase, when he nominates the two concepts that are to be the topic of his discourse: "*cognizione*" and "*linguaggio*". As he nominates the two concepts, first his left hand and then his right hand are positioned on either side of his body in space. The hands act as if two entities are being held for view, each one on either side. Here, by the way the speaker uses his hands, it is as if he is holding up two different objects. Because, as he brings his left hand into position as he says "*cognizione*" and as he brings his right hand into position he says "*linguaggio*", we are inclined to interpret this to mean that the two virtual entities being held up are labelled "*cognizione*" and "*linguaggio*", respectively. For this reason we are inclined to say that the speaker is treating "cognition" and "language" as if they are two separate concrete objects which can be held in the hand.

In his next phrase the speaker goes on to say something about the relationship between these two concepts and, as we seen, having set them up with his hand action as if he has two objects which represent them, he now moves his two hands in relation to one another in such a way as to provide a kind of dynamic diagram of the relationship between the concepts as he describes this in words. He begins phrase 4 by saying "*sono aspetti*" – here he is referring to both concepts at once,

and it is notable that he lowers both his hands in unison over “*sono*” and raises them and separates them in space, again in unison as he says “*aspetti*”. Talking of both entities together, as he does in what he says, the movement in unison of his two hands seems to reflect this. The two objects are being treated together, as they are being referred to together. Then, as he says “*dissociabili tra loro*”, he brings his two hands together in space and then rapidly separates them – as if showing how, if these two entities are placed in the same space, they will not remain there but move away from one another into separate spaces, in this way presenting in the form of dynamic spatial action, their non-association.

What this example illustrates, then, is that here the hands are used in actions that are like holding objects or moving objects about in space in the context of a discourse which describes abstract relationships between conceptual entities. This is a very common use of gesture and it may be compared to the illustration of abstract concepts with drawings on a blackboard or, as sometimes happens in classrooms, the use of coloured balls or other objects to stand for geometrical or other mathematical concepts.

### 2.3. Example 3

The third example I would now like to present is taken from a video-recording I made with some students of mine in Naples some years ago in which ordinary Neapolitans were asked to talk about being Neapolitan or to talk about what are, for them, the important characteristics of Neapolitan culture. Here one of the students, Massimo, asks someone to give a typical recipe. He says:

- (3) *Allora/sulla cucina/ci può dare una ricetta/tipicamente napoletana/  
descrivere cioè/dall'alla zeta/per quattro persone  
well/about cooking/can you give a recipe/typically Neapolitan/  
describe, that is/from A to Z/for four persons*

We may divide Massimo's discourse into seven tone units or intonation phrases. In association with each one of these successive tone units we see a distinctive pattern of action in the hands. Already, at the level of the higher level units of production - that of speech phrases and phrases of gestural action - we can see a nice correspondence, a further illustration of the coordinate production of spoken and manual expressions. If we consider, now, the forms of manual action that Massimo produces here we see that what most of his expressions do is to provide a kind of enactment of the *pragmatic effect* of the corresponding unit of speech.

Thus, as he says “*sulla cucina*” (nominating the topic), he brings his hands together in a bunched fashion, in a form of action not unrelated to what people often do when they nominate a topic (Kendon 2004:230). Then, as he says “*puoi darmi una ricetta*” (here politely asking for something), he puts his hands together palm to palm in an expression that is very well known, often used (in Neapolitan culture) when people make polite requests, although also in certain other request-

ing contexts, too (Kendon 1995). He combines this with a bowing forward of the body, thus “doing” the bodily ritual of “polite request” even as his words make a request also. However, in his words he is able to nominate what he is requesting, which is a “*ricetta*”.

He then says “*tipicamente napoletana*”. By saying this, he limits the range of recipes he wants described. As he does so he uses a two handed expression that we might call “field delimiting” or “topic delimiting” - a form of action often used when people specify some field for consideration, as it were delimiting the ‘space’ within which the discourse is to be confined or focused (Kendon 2004:104). He then continues by saying “*descrivere cioè*”, thus making much more specific the task he wants his interlocutor to undertake. Here he uses a hand shape in which tip of index finger bends over to touch tip of thumb (sometimes called a “ring” hand shape) which is often used in contexts where the speaker is indicating that what he is saying is a highly specific focus within the wider field that he has established in his discourse (see Kendon 2004:342). He continues with “*dall’a alla zeta*”. Here he moves his hand laterally in a horizontal fashion. Perhaps, with his hand moving through space in this fashion, he makes manifest in the form of a spatial movement the idea of passing across a range of locations. He makes visibly concrete the spatial metaphor that is implicit in the expression “*dall’a alla zeta*”. Finally he says “*per quattro persone*”, holding his hand up displaying four fingers as he does so. Here he presents a visible kinesic sign for the numeral “*quattro*”.

Once again, in this example, we see how the manual actions are coordinated with the speech – the phrasal units of speech and the phrasal units of gesture are produced together and thus must be seen as integral components of the utterance action.

Second, as I have suggested in my brief analysis, the forms of manual action or gestures in this example are forms that are often used and, by comparing the contexts in which we see them used, we can identify rather general meanings for them. This leads me to say, for instance, that we have actions that are done in contexts of ‘topic nomination’, ‘topic delimiting’, that are done to mark a certain kind of request, and that are used in relation to a speaker’s attempt to *specify* something. As I have said, these are forms that are, so to speak, kinesic versions of *doing* actions of this kind - doing the acts that the speaker is doing also with his speech and this is why we call them ‘pragmatic’.

I should add here that what I have identified as ‘recurrent forms’ which permit a ‘context of use analysis’ have been identified mainly among speakers in Italy, mainly in the Neapolitan area. Some of these may well be characteristic of Neapolitans - the bowing forward with the hands held palm-to-palm is not used much in England, for example. Others have a much wider distribution - the ‘topic delimiting’ gesture, for example, appears to be quite widespread.

Third, if we accept these meanings we can see how, as we saw in the two previous examples, the ‘semantic level’ at which these actions relate to the semantics of the speech is variable. I have suggested how the first four appear to give visible

actions that relate to the *pragmatic actions* of what the speaker is saying. The last two, however, employ action to present aspects of the referential content of what he is saying. A horizontal movement in connection with a verbal expression which refers to the idea of a succession or a series of actions, and a presentation of the hand shaped in a certain way which is a conventional kinesic representation of the numeral “*quattro*”.

### 3. The Lessons of the Three Examples

From these three examples, then, we can see, first, how delicately intimate is the coordinative relationship between these rather large manual movements that we perceive as “gestures” and the speech that is produced with them. Gestures and speech are produced *together* as products of a single, unitary, project of action. This is why I speak of the *gesture-speech-ensemble*.<sup>2</sup> Yet, if, in this case, we must admit these movements as partners with spoken elements in the utterances in question, we nevertheless must see that the roles they serve in these utterances is quite varied.

In the first example, we interpreted the movements as serving as contributions to the descriptions of the objects the speaker was talking about. We suggested that they entered into the construction of her utterance at the level of its propositional meaning and proposed that, in this case, we could incorporate the gestures into her utterance linguistically, if we accepted the idea that speakers, by using kinesic as well as verbal resources, could engage in ‘simultaneous construction’. In the second example, on the other hand, we saw kinesic actions that worked in a much more ‘parallel’ manner. The upheld hands posed as if holding objects and the way these were subsequently moved in relation to one another were interpreted as presenting in kinesic form a *diagram* of the objects and their relationships that the speaker was talking about. In the final example we illustrated the so-called ‘pragmatic’ uses of gesture. Here, the kinesic expressions served to mark or express, as schematic forms of action, the illocutionary acts that the speaker was accomplishing with his utterance.

In the manual actions that speakers employ when speaking, thus, we see that they can do many different things and that the way speakers’ manual actions relate, semantically, to the semantics of the speech can also be very different. A careful examination of even such commonplace examples as I have presented here, shows that the speaker is doing something with gestures that is intimately intertwined with what is being done with words and the gestures are employed with a diversity of purpose as the utterance unfolds.

---

<sup>2</sup> The recognition that gesture and speech are produced together as components of a single plan of action was developed in Kendon (1972, 1980) and elaborated extensively by McNeill (1992). The term “gesture-speech ensemble” is introduced in Kendon (2004).

#### **4. Concluding Remarks**

We have seen, then, that what these examples suggest is that engaging in the act of utterance - that is, in engaging in the production of a unit or units of action that operate as 'contributions' in situations where talk is the focused activity - 'speakers' (as we call them) draw upon a variety of devices or resources and construct their utterances as a complex orchestration of verbal spoken actions and kinesic or gestural actions. It seems clear that, from the point of view of the way in which 'linguaging' is done (if I may be permitted such a verb), we must see it as indeed a *multimodal* ensemble of actions.

However, I think we still remain too much dominated by the idea that verbal language is an autonomous system which does not *need* these other forms of expression. It can get along quite well without them. And indeed, so it can. We especially have this idea because of the power of written language. But the question is, not *can* verbal language get along without other forms of expression, but *does* it and, if it does, when does it do so? That people routinely employ kinesic devices of various kinds as they actually construct utterances, suggests, I think, that from the producer's point of view verbal expression is often not enough and other devices must be employed. The model of verbal language that is used in theoretical thinking, however, is, to a large extent, a model based on an *abstraction* from the complex of actions involved in utterance production. As already mentioned, this is abstraction that writing has made possible and which, indeed, it nurtures. But this model, let us call this the autonomy model, does not give a satisfactory account of how people actually produce utterances much of the time. It has been the intent of my examples to suggest this.

Now I would like to add the comment that it is this autonomy model of language that has been the guiding model for the description not only of verbal language but for the description of signed languages, also. When, in 1960, Stokoe published his famous *Sign Language Structure* he set out to demonstrate that the model of language that was then current in American structural linguistics, which he had got largely from the work of Trager and Smith, could be applied to the visual communication system of the deaf that he had encountered at Gallaudet College (as it then was). Following this demonstration, which played an important role in the process by which such visual communication systems got recognition as being indeed *languages* (with all the prestige that comes when that label is seen to be deserved), it became a matter of high priority by many, in those early days of sign language research, to demonstrate in detail just how well the autonomy model of language could be employed to describe sign language - to such an extent that various aspects of sign language that did not fit well with this, such as 'iconicity' and the 'gesture-like' aspects of space use and the use of so-called 'classifiers', were either played down in importance or were incorporated into systematic linguistic descriptions in ways that often seemed highly cumbersome and forced (Wilcox 2004).

Gradually, however, students of sign languages have come to recognize that the autonomy model of language derived from the analysis of verbal language

does not fit, and a different model of language will have to be developed to deal with sign language. This is because, in the actual production of signed utterances expressive devices must be employed which don't fit the spoken language model. I refer here, for example, to Scott Liddell's (2003) work on the structuring of space in signed discourse and his argument that the way in which pronominal reference is accomplished is through processes which involve 'mapping' real and virtual spaces onto one another, rather than through the employment of an ever more complicated inventory of discrete 'morphemes'. There has also developed, in recent years, a debate about whether signers 'gesture' or not - but here again, it seems to me, that this discussion is based upon the idea that there is something that is called a 'language' - autonomous and separate - and this accounts for certain things, but what lies outside it must be 'gesture'. And yet it is becoming clear, both through Liddell's work, as I have already mentioned, but also through the work of others, that modes of expression used freely and commonly in sign language discourse which yet don't fit the autonomy model, are actually an integral part of the way signers construct utterances (see discussions in Emmorey 2003, and see Cuxac and Sallandre 2007, Fusellier-Souza 2006, Pietandrea and Russo 2007, among many others). To separate out these expressions that don't fit as "gesture" or to worry about whether they are "gesture" rather than "sign" seems to be a product of the persistence of the autonomy model and our failure to develop a unified model of utterance production that could incorporate them in an integrated way.

As I have suggested (Kendon 2008), I think it would be useful if we were to abandon altogether the use of both the term "gesture" and the term "sign" and try, instead, to develop a comparative semiotics of the utterance uses of kinesic action. If we do this we can, in the case of signed languages, see how, in signing, a wide range of semiotic devices are employed - all the way from Cuxac's "highly iconic structures" and the flexible way in which space is employed, to stable, codified or "frozen" forms which may be even be analyzable in terms of something analogous to a phonology in spoken language.

However, what we also see is that, if we look at people who use speech, and observe them as they construct utterances - as I have tried to illustrate with my examples today - we see how also they use a range of semiotic devices and, in the kinesic modality this range goes all the way from forms that appear to be highly codified, that work, as I have suggested, as devices that label conceptual categories (just like words do, or so-called "frozen" signs in sign language do) to forms that are quite like the "highly iconic structures" that Cuxac describes in his work on French Sign Language.

So if we compare instances of "languaging" - whether this is done entirely in the kinesic modality (as is the case in signing) or in a combination of the spoken and the kinesic - we see how in both there is a flexible use of a range of semiotic devices. If we want to compare signing with speaking we must compare signing with speaking as it is *performed*. Sign language does not have the benefit of millennia of a writing system and it can exist, thus, *only* in a performed manner.

## *Kinesic Components of Multimodal Utterances*

Speakers can, and do, often organize their spoken utterances in relation to the forms and patterns developed as a consequence of language in its written form, but they do not always do so. And when we look at speakers as they are actively producing utterances we see that they, also, make use of a wide range of strategies.

Scott Liddell (2003) concludes from his analysis of American Sign Language in which he sees what he calls the “gestural” as being an integral part of its grammar, that this might lead us to develop a new way of thinking about language in general. What I hope I have suggested here is that, through the little examples I have provided, this new way of thinking could be highly useful, for it might provide us the basis for the *integrating* model that we need, if we are to have a good theory of *utterance* production and not just a theory of the construction of discourse in only one modality.

### **References**

- Austin, Gilbert. 1806. *Chironomia or, a treatise on rhetorical delivery*. M. Margaret Robb and L. Thonssen, eds. Carbondale and Edwardville: Southern Illinois University Press, 1966.
- Barnett, Dene. 1987. *The art of gesture: The practices and principles of 18th century acting*. Heidelberg: C. Winter.
- Beattie, Geoffrey, and Heather Shovelton. 1999. Do iconic hand gestures really contribute anything to the semantic information conveyed by speech? An experimental investigation. *Semiotica* 123:1-30.
- Beattie, Geoffrey, and Heather Shovelton. 2001. An experimental investigation of the role of different types of iconic gesture in communication: A semantic feature approach. *Gesture* 1:129-49.
- Beattie, Geoffrey, and Heather Shovelton. 2002. An experimental investigation of some properties of individual iconic gestures that affect their communicative power. *British Journal of Psychology* 93 (2):179-72.
- Butler, H. E., trans. 1922. *The institutio oratoria of Quintilian*, vol. 4. The Loeb Classical Library. London: William Heinemann.
- Cienki, A., and C. Müller, eds. 2008. *Metaphor and gesture*. Amsterdam and Philadelphia: John Benjamins.
- Cuxac, Christian, and Marie-Anne Sallandre. 2007. Iconicity and arbitrariness in French Sign Language: Highly iconic structures, degenerated iconicity and diagrammatic iconicity. In E. Pizzuto, P. Pietandrea, and R. Simone, eds., *Verbal and signed languages: Comparing structures, concepts and methodologies*, 13-33. Berlin: Mouton de Gruyter.
- De Jorio, Andrea. 2000. *Gesture in Naples and gesture in classical antiquity. A translation of La mimica degli antichi investigata nel gestire napoletano*. A. Kendon, trans. and ed. Bloomington: Indiana University Press. (Orig. pub. 1832.)

- Emmorey, Karen, ed. 2003. *Perspectives on classifier constructions in sign languages*. Mahwah, NJ: Lawrence Erlbaum.
- Fusellier-Souza, Ivani. 2006. Emergence and development of signed languages: From a semiogenetic point of view. *Sign Language Studies* 7(1):30-56.
- Goldin-Meadow, Susan. 2003. *Hearing gesture: How our hands help us think*. Cambridge, MA: Harvard University Press.
- Harris, Roy. 1987. *The language machine*. Ithaca, NY: Cornell University Press.
- Kendon, Adam. 1972. Some relationships between body motion and speech. An analysis of an example. In A. Siegman and B. Pope, eds., *Studies in dyadic communication*, 177-210. Elmsford, NY: Pergamon Press.
- Kendon, Adam. 1980. Gesticulation and speech: Two aspects of the process of utterance. In M. R. Key, ed. *The relationship of verbal and nonverbal communication*, 207-227. The Hague: Mouton and Co.
- Kendon, Adam. 1981. Geography of gesture. *Semiotica* 37:129-163.
- Kendon, Adam. 1984. Did gesture have the happiness to escape the curse at the confusion of Babel? In A. Wolfgang, ed., *Nonverbal behavior: Perspectives, applications, intercultural insights*, 75-114. Toronto: C. J. Hogrefe, Inc.
- Kendon, Adam. 1995. Gestures as illocutionary and discourse structure markers in southern Italian conversation. *Journal of Pragmatics* 23:247-279.
- Kendon, Adam. 2004. *Gesture: Visible action as utterance*. Cambridge: Cambridge University Press.
- Kendon, Adam. 2008. Some reflections on 'gesture' and 'sign'. *Gesture* 8(3):348-336.
- Kita, Sotaro. 2000. How representational gestures help speaking. In D. McNeill, ed., *Language and gesture*, 162-185. Cambridge: Cambridge University Press.
- Krauss, Robert M., Yihsiu Chen, and Rebecca F. Gottesman. 2000. Lexical gestures and lexical access: A process model. In D. McNeill, ed., *Language and gesture*, 261-283. Cambridge: Cambridge University Press.
- Liddell, Scott K. 2003. *Grammar, gesture and meaning in American Sign Language*. Cambridge: Cambridge University Press.
- Linell, Per. 2005. *The written language bias in linguistics: Its nature, origins and transformations*. London: Routledge.
- McNeill, David. 1992. *Hand and mind*. Chicago: Chicago University Press.
- McNeill, David. 2005. *Gesture and thought*. Chicago: Chicago University Press.
- Melinger, Alissa, and Willem J.M. Levelt. 2004. Gesture and the communicative intention of the speaker. *Gesture* 4(2):119-41.
- Payrató, Lluís. 1993. A pragmatic view on autonomous gestures: A first repertoire of Catalan emblems. *Journal of Pragmatics* 20:193-216.
- Pietandrea, Paola, and Tommaso Russo. 2007. Diagrammatic and imagic hypicons in signed and verbal languages. In E. Pizzuto, P. Pietandrea and R. Simone, eds., *Verbal and signed languages: Comparing structures, constructs and methodologies*, 35-56. Berlin: Mouton de Gruyter.

*Kinesic Components of Multimodal Utterances*

- Slobin, Daniel I. 1987. Thinking for speaking. In J. Aske, N. Beery, L. Michaelis, and H. Filip, eds., *Proceedings of the 13<sup>th</sup> Annual Meeting of the Berkeley Linguistics Society*, 435-444. Berkeley, CA: Berkeley Linguistics Society
- Stokoe, William C. 1960. Sign language structure: An outline of the visual communication systems of the American Deaf. *Studies in Linguistics Occasional Papers* 8:1-78. Buffalo, NY: Department of Anthropology and Linguistics, University of Buffalo.
- Streeck, Jurgen. 2008. Depicting by gesture. *Gesture* 8(3):285-301.
- Streeck, Jurgen. 2009. *Gesturecraft: The manu-facturing of meaning*. Amsterdam and Philadelphia: John Benjamins.
- Vermeerbergen, Myriam, Lorraine Leeson, and Onno Crasborn, eds. 2007. *Simultaneity in signed languages: Form and function*. Amsterdam: John Benjamins.
- Wilcox, Sherman E. 2004. Hands and bodies, minds and souls: What can signed languages tell us about the origin of signs? In M. Alac and P. Violi, eds., *In the beginning: Origins of semiosis*, 137-67. Turnhout, Belgium: Brepols.
- Woll, Bencie. 2007. Perspectives on linearity and simultaneity. In M. Vermeerbergen, L. Leeson and O. Crasborn, eds., *Simultaneity in signed languages*, 337-344. Amsterdam: John Benjamins.

Adam Kendon  
43 West Walnut Lane  
Philadelphia, PA 19144  
USA

adamk@dca.net