

## 8 CONCLUSION

---

In battle there are only the normal and extraordinary forces, but their combinations are limitless.

Sun Tzu

### 8.1 The anatomy of meaning: summary remarks

In solving the ever-present puzzle of figuring out what others are trying to say, our evidence comes in chunks: composite utterances built from multiple signs of multiple types. These composites are produced by people on trajectories of collaborative social activity. As communicative behaviours, they are strategic, context-embedded efforts to make social goals recognizable. The essential claim of this book is that if we are to understand how people interpret such efforts, our primary unit of analysis must be the utterance or move, the single increment in a sequence of social interaction. Component signs will only make sense in terms of how they contribute to the function of the move as a whole.

This book has focused on moves built from speech-with-gesture as a sample domain for exploring the anatomy of meaning. But the analytic requirement to think in terms of composite utterances is not unique to speech-with-gesture. Because all utterances are composite in kind, our findings on speech-with-gesture should help us to understand meaning more generally. This is because research on the comprehension of speech-with-gesture is a sub-field of a more general pursuit: to learn how it is that interpreters understand token contributions to situated sequences of social interaction (cf. Schegloff 1968, Goffman 1981). Chapter 1 asked how multiple signs are brought together in unified interpretations, framing the issue in terms of semiotic function of a composite's distinct components (see Figure 1.7). A broad distinction was made between conventional meaning and non-conventional meaning, where these two may be joined by indexical mechanisms of various kinds. Think of a painting like Bouguereau's *Wave* (Figure 1.3): a title (words, conventional) is taken to belong with an image (an arrangement of paint, non-conventional) via indexical links (spatial co-placement on a gallery wall, putative source in

a single creator and single act of creation). The speech-with-gesture composites discussed in this book can be analysed in the same way. When a man says *Make it steep like this* with eye gaze fixed on his arm held at an angle (see Figure 1.4), the conventional signs of his speech are joined to the non-conventional sign of his arm gesture by means of indexical devices including temporal co-placement, source in a single producer, eye gaze, and the symbolic indexical expression *like this*. Or when a man says *Make it fluted at the mouth* with eye gaze fixed on his symmetrical 'fluted' hands (see Figure 1.5), conventional word and non-conventional hand are joined indexically by temporal co-placement, source in a single producer, and eye gaze. In these illustrative gesture cases (see Part II of this book), hand movements constitute the non-conventional 'image' component of the utterance. By contrast, in deictic gesture or pointing cases (see Part I), hand movement provides the indexical link between words and some image or thing in the world, such as buffaloes walking by (see Figure 2.9), or diagrams in ink or mid-air (see Chapters 6 and 7).

This semiotic framework permits systematic comparison of speech-with-gesture moves to other species of composite utterance. An important case is sign language of the Deaf. There is considerable controversy as to how, if at all, gesture and sign language are to be compared (cf. Emmorey and Reilly 1995). The present account makes it clear that the visible components of a sign language utterance cannot be compared directly to the visible hand movements that accompany speech, nor to mere speech alone (with visible hand movements subtracted), but may only be properly compared to the entire speech-with-gesture composite (cf. Liddell 2003). The unit of comparison must be the move. By the analysis advanced here, different components of a sign language move will have different semiotic functions, in the sense just discussed: conventional signs with non-conventional signs, linked indexically. Take the example of sign language 'classifier constructions' or 'depicting verbs' (Liddell 2003: 261ff.). In a typical construction of this kind, a single articulator (the hand) will be the vehicle for both a conventional sign component (a conventionalized hand shape such as the ASL 'vehicle classifier') and a non-conventional sign component (some path of movement, often relative to a contextually established set of token spatial referents), where linking indexical mechanisms such as spatio-temporal co-placement and source in a single creator are maximized through instantiation in a single sign vehicle, i.e. one and the same hand.

Another domain in which a general composite utterance analysis should fit is in linguistic research on syntax. Syntactic constructions, too, are made up of multiple signs, where these are mostly the conventional signs of morphemes and constructions (though note of course that many grammatical morphemes are symbolic indexicals). An increasingly popular view of syntax

takes lexical items (words, morphemes) and grammatical configurations (constructions) to be instances of the same thing: linguistic signs (Langacker 1987, Goldberg 1995, Croft 2001). From this viewpoint, interpretation of speech-only utterances is just as for speech-with-gesture. It means dealing with multiple, simultaneously occurring signs (e.g. *That guy* may be both noun phrase and sentential subject), and looking to determine an overall target meaning for the communicative move that these signs are converging to signify. A difference is that while semantic relations within grammatical structures are often narrowly determined by conventions like word order, speech-with-gesture composites appear to involve mere co-occurrence of signs, with no further formal instruction for interpreters as to how their meanings are to be unified. Because of this extreme under-determination of semiotic relation between, say, a gesture and its accompanying speech, many researchers conclude that there are no systematic combinatorics in speech-with-gesture. But speech-with-gesture composites are merely a limiting case in the range of ways that signs combine: all an interpreter knows is that these signs are to be taken together, but there may be no conventionally coded constraints on **how**. Such under-determination is not unique to gesture. In language, too, we find minimal interpretive constraints on syntactic combinations within the clause, as documented for example by Gil (2005) for extreme isolating grammar found in some spoken languages. And beyond the clause level, such under-determined relations are the standard fabric of textual cohesion (Halliday and Hasan 1976).

In sum, to understand the process of interpreting any type of composite utterance, we do not begin with components like noun, rising intonation, or pointing gesture. We begin instead with the notion of a whole utterance, a complete unit of social action which always has multiple components, which is always embedded in a sequential context (simultaneously an effect of something prior and a cause of something next), and whose interpretation always draws on both conventional and non-conventional signs, joined indexically as wholes.

## **8.2 Semiotic unification – towards a rational, heuristic-based model**

This book is a case study in the general problem of semiotic unification: how is it that an interpreter takes multiple signs to stand together for a coherent whole meaning? To treat the problem in the most general terms possible, we can consider utterance comprehension as a species of decision-making. The interpreter's task is to decide what someone is trying to say (or what someone is trying to do by what they are saying). Assume that interpreters in social

interaction follow the same three-step strategy they follow in other decision-making domains (Gigerenzer *et al.* 1999):

- 1 **lock-on** to a problem to be solved (i.e. identify a target decision to be made and instigate a search);
- 2 **narrow the search** as far as possible;
- 3 **lock-off** by making a decision appropriate for current purposes, thereby stopping the search.

In the domain of interpreting others' contributions to conversation (typically, composites of speech and visible behaviour), these phases in a process of rational decision-making can be thought of as follows:

- 1 lock-on by recognizing that a swatch of behaviour has a communicative intention, and instigate a search for the informative intention behind it;
- 2 use sign filtration to constrain the search for relevant signs within that behaviour at lowest cost within a search space;
- 3 lock-off by making a currently appropriate interpretation of the utterance which constitutes a satisfactory stop to the search.

An adequate account of semiotic unification in the interpretation of addressed communicative utterances will have to map out how the three phases of rational interpretation are navigated in the enchronic environment of conversation. The following paragraphs offer some initial thoughts in this direction, based on considerations raised in Chapter 1, and explored in the empirical studies in Chapters 2–7.

### 8.2.1 *On-switch*

What causes an interpreter to lock on to a problem of communicative interpretation in the first place? The answer is any sign of communicative intention, i.e. any indication that a person is carrying out a controlled, addressed behaviour in order that the behaviour be taken by another person to be a sign of saying something (Grice 1957). A simple heuristic for recognizing communicative intention is to identify a behaviour whose very existence is already dedicated to that function – for example, any conventional sign such as the words and constructions of language. Other heuristics include simple cues associated with the everyday use of such signs. For example, if someone is looking straight at your face when they carry out some action, it is more likely than not that they are inviting you to attend to that action (for example they may be doing the action in order to demonstrate something to you; Csibra and Gergely 2006, Gergely and Csibra 2006).

Further heuristics for recognizing when to lock on to a behaviour and begin a search for meaning include general abductive mechanisms of rational interpretation. For example, if a person's action is done in an unusual or otherwise marked manner, or in a way that cannot be explained for obvious practical reasons, then an interpreter may conclude that this person must be doing this marked action for some special purpose, and lock on to this as a puzzle which deserves to be solved (Grice 1975). Prelinguistic infants do this when seeing adults carry out actions in unconventional ways such as turning on a light using the head rather than the hands (Gergely *et al.* 2002). Or suppose I'm telling you about someone I just met and I say, suddenly speaking at noticeably increased volume: HE SPEAKS REALLY LOUD. Your sign filter will take the increased volume to mean something, and by regarding the words together with the non-conventional sign of suddenly increased volume (indexically cemented by co-occurrence in a single sign vehicle: the speech stream), the composite utterance will readily be taken to both describe and illustrate how he speaks.<sup>1</sup>

The types of composite utterance described in this book possess most or all of the features which these lock-on heuristics would exploit. They are thus straightforwardly taken by interpreters to be communicative problems in need of interpretive solutions. Of course, these utterances are not mere passing puzzles of interest. Interlocutors are highly motivated to deal adequately with others' moves due to the social consequences inherent in each increment to a sequence of social interaction (Schegloff 1968, Heritage and Atkinson 1984, Enfield 2006), and the morally-grounded joint commitment which social interaction entails (Clark 2006).

### 8.2.2 Search: sign filtration

Once an interpreter has locked on to a person's actions as communicatively intended, they need to know which are the signs that they should attend to, and which are not, in order to assess a signer's (putative) informative intentions – i.e. the content of what they're trying to say.<sup>2</sup> As Kendon points out (e.g. 1986, 2004), interpreters display finely tuned differentiation of attention when faced with speech and visible behaviour together. Again, conventional signs like linguistic items are straightforwardly recognized. Their very *raison d'être* is to convey communicative and informative intentions

<sup>1</sup> A variant would be HE SPEAKS **LIKE THIS**, using the symbolic indexical 'like this' to explicitly send you on a search for how it is that he speaks – i.e. supplying only an illustration, not a description (Engle 1998, Clark 1996).

<sup>2</sup> It does not matter whether a speaker actually has the psychological state corresponding to what is putatively conveyed by the signs. For instance, applause may be insincere, but it always **claims** to praise.

in the service of social action. (The same goes for pointing gestures, since as symbolic indexicals they incorporate conventional meaning.)

Now when a person produces conventional signs, they also produce a stream of other available information in the form of non-conventional signs. Consider the hand gestures described in Part II of this book. These, too, provide no significant sign filtration problem for interpreters, since they are gazed at, pointed to, spoken about, and positioned squarely in the attentional field of their intended addressees (cf. Clark 1996, Goodwin 2000a). In addition, these hand movements are straightforwardly connected to what the speaker is saying, in timing and in speakers' dynamic investment in the communicative activity (Levy and Fowler 2000). As such they are more likely to be taken as signs of a speaker's informative intention than, say, a random scratch of the head. Also, it seems clear that practices of pointing, tracing, diagramming, and modelling with the hands as exemplified in this book are themselves directly recognizable as conventional communicative practices (Kendon 2004, Wilkins 2006).

### 8.2.3 *Off-switch*

What determines that our ongoing interpretation of an ensemble of signs is sufficient for current purposes, such that no further interpretation is required?<sup>3</sup> This is the sixty-four-thousand-dollar question. Few have given it serious attention (a notable exception being Sperber and Wilson 1995 [1986]). One thing for sure is that interpreters do not routinely make an exhaustive analysis of all signs presented in a given move (Sanford and Sturt 2002). Lock-off cannot be contingent upon a complete reading of all signs, conventional or otherwise, in a composite utterance. If it were, we would get stuck hanging on every potentially meaningful element of someone's action, unable to move forward.

If an interpreter is applying fast and frugal heuristics (Gigerenzer *et al.* 1999), the point is to lock-off on an interpretation which is good enough for current purposes. The objective is an optimal trade-off between minimizing effort and maximizing yield. Perceived communicative needs of the context will determine what is an optimal degree of attention paid to available signs and their conceivable meanings. Now if interpretations can be functionally adequate on the basis of less-than-exhaustive analyses of the available signs, this means that there are components of utterances that are dispensable. Can any generalizations be made as to what kinds of utterance components are more dispensable than others? Research is needed here.

<sup>3</sup> Lock-off by interpreters is a correlate of McNeill's 'stop order' for producers (McNeill 2005: 18).

To understand how lock-off comes about, it is worth being mindful of the fact that interpreters do not work (purely) from the ground up. From the start, an interpreter's search is for something which **gives rise to** a producer's signs, that is, the intention which those signs should make recognizable. Upon recognition that a signer is saying something (i.e. recognition that there is a communicative intention) an interpreter searches for an informative intention. This will not be found by merely decoding and summing up the available signs, as if they were a carbon copy of the speaker's token meaning. That meaning can only be **suggested by** the signs. How we get from sheer form to speaker-meaning remains unknown, but the mechanism is bound to be simpler than it appears. An over-arching principle for lock-off in discovering what others mean (in the spirit of Sperber and Wilson 1995 [1986]) might be as follows:<sup>4</sup> when you see a set of signs, and see them **as** signs, take them to stand for the first object that they could simultaneously stand for.

#### 8.2.4 *A heuristics-based approach*

With an eye to further directions in research on composite utterances, I have hinted at some elements of a heuristics-based approach to analysing how moves as social actions are understood on the basis of formal patterns of behaviour in interaction. In Chapter 1, I suggested a preliminary list of interpretive triggers and heuristics, including a convention heuristic, an orientation heuristic, a contextual association heuristic, a unified utterance-meaning heuristic, and an agency heuristic. We can also expect to rely heavily on a relevance heuristic, inherent in an enchronic model of meaning and its emphasis not just on moves but on relations between moves. A central task for subsequent research is to discover, describe, and test the full set of heuristics which make tractable the mind-boggling yet seemingly effortless task of figuring out what others are trying to say.

### 8.3 **Concluding remark: speech, gesture, and meaning**

Research on speech-with-gesture yields ample motivation to question the standard focus in mainstream linguistics on competence and static representations of meaning (as opposed to performance and dynamic processes of meaning; see McNeill 2005: 64ff., Wilkins 2006: 140–141). There is a need for due attention to meaning at a context-situated token level (a stance preferred by many functionalist linguists, linguistic anthropologists, conversation

<sup>4</sup> With thanks to Paul Kockelman for this pithy phrasing of a thought that once took me an hour to express. Note that 'object' here is meant in the technical sense of Peirce (1955); i.e. whatever a sign stands for (not necessarily a physical object).

analysts, and some gesture researchers), rather than privileging the analysis of abstract types (preferred by a Saussurean mainstream). Speech-with-gesture composites quickly make this need apparent, because they force us to examine singularities, i.e. semiotic structures that are tokens but not tokens-of-types. These singularities include non-conventional gestures as utterance components, as well as the overall utterances themselves, each a unique combination of signs. This is why, for instance, Kendon writes of speech-with-gesture composites that 'it is only by studying them as they appear within situations of interaction that we can understand how they serve in communication' (2004: 47–48). Here is the key point: what Kendon writes is already true of speech whether it is accompanied by gesture or not (e.g. Hanks 1990, 1996 among many others). It is just that these lessons are not taught in mainstream linguistics. Speech-with-gesture teaches us to address meaning at the token-level, something we should be doing anyway.

The point is reminiscent of Theodore Roszak's (1977) argument that the changes in social practice necessary to improve our health and well-being locally are the same changes necessary to avert environmental disaster globally. Or as Roszak put it: 'The needs of the person are the needs of the planet.' In the same spirit, the analytic stance that speech-with-gesture demands – i.e. to treat moves as dynamic, motivated, concrete, and context-bound – is the same stance we need for the proper treatment of communicative moves generally, including the subject matter of linguistics, anthropology, and other branches of semiotics. In other words: the needs of research on gesture are the needs of research on meaning.