7 A modality-free notion of gesture and how it can help us with the morpheme vs. gesture question in sign language linguistics (Or at least give us some criteria to work with)

Arika Okrent

7.1 Liddell's proposal that there are gestures in agreement verbs

Forty years of research on signed languages has revealed the unquestionable fact that signers construct their utterances in a structured way from units that are defined within a language system. They do not pantomime or "draw pictures in the air." But does this mean that every aspect of a signed articulation should have the same status as a linguistic unit?

A proposal by Liddell (1995; 1996; Liddell and Metzger 1998) has brought the issue of the linguistic status of certain parts of American Sign Language (ASL) utterances to the fore. He proposes that agreement verbs are not verbs simultaneously articulated with agreement morphemes, but verbs simultaneously articulated with pointing gestures. Agreement verbs are verbs that move to locations in signing space associated with particular referents in the discourse. A signer may establish a man on the left side at location *x* and a woman on the right side at location *y*. Then, to sign 'He asks her,' the signer moves the lexical sign ASK from location *x* to location *y*. The locations in these constructions have been analyzed as agreement morphemes (Fischer and Gough 1978; Klima and Bellugi 1979; Padden 1988; Liddell and Johnson 1989; Lillo-Martin and Klima 1990; Aarons *et al.* 1992) that combine with the lexical verb to form a multimorphemic sign *x*ASK*y*. However, according to Liddell's claim, when a signer produces the utterance 'He asks her,' he or she does not combine the specifications of the sign ASK with morphemic location features, but rather, he or she points at cognitively meaningful locations in space with the sign ASK. This challenges the claim that the locus aspect of these utterances is morphemic, or even phonological.

Liddell argues for a gesture account partly on the basis of the impossibility of specifying the form of the agreement morpheme. There are an infinite number of possible locations that can have referential value. The signing space cannot be divided into 10 or 20 discrete, contrastive locations; there are rather as many locations possible as there are discernible points in space. The locations used

are, in situations where the participants are present, determined by the position of real people in real space. If the man and the woman being discussed were present in the above example, the verb would move from the real location of the man to the real location of the woman. The form of the agreement morpheme can only be described as "the place where the thing I'm talking about is located."[1] The location of a person in space is a fact about the world independently of anybody's language and should not be considered linguistic.

7.2   Objections to the proposal[2]

Liddell's proposal has been controversial. One of the main objections to the notion of verb agreement being gestural has been that the pointing[3] in agreement verbs is restricted by language-internal considerations.

- **Some verbs can do it and some verbs cannot:** The verb LOVE cannot point toward its object, but the verb ASK can. One must know the language to know which verbs can and cannot point.
- **Verbs which point, must point, and must do so correctly:** The verb ASK must point to the subject and then the object while the verb INVITE must point to the object and then the subject. A verb that points must point in a sentence where its arguments are explicitly located: *HEx ASK(no pointing) HERy. One must know the language to know the proper pointing behaviors for verbs.
- **Different sign languages do it in different ways:** In the Sign Language of the Netherlands, the verb itself is articulated in neutral space, followed by an auxiliary element that points (Bos 1996). One must know the language to know whether it is the verb or an auxiliary that points.

The implication behind these objections is that, even if it is true that the phonological form of the agreement morpheme is unspecified, the pointing in general is restricted on the basis of language internal considerations, and what is restricted on the basis of language internal considerations cannot be gesture.[4]

---

[1] For the case where the actual referents are not present, Liddell (1995; 1996) argues that they are made present through the grounding of mental spaces, something that is also done by nonsigners.

[2] The authors cited for the objections did not themselves frame any of their works as arguments against the gesture account of agreement.

[3] "Pointing" generally refers to an action performed with an index finger. Here "pointing" refers to the same action, but without necessarily involving the use of an index finger. It is "pointing" performed with whatever handshape the verb is specified for.

[4] There are also objections based on psycholinguistic studies of language development (Newport and Meier 1985; Meier 1987; Petitto 1987) and brain damage (Poizner *et al.* 1987). I do not address these objections in this chapter, but I do not believe the results of these studies necessarily rule out the gesture account. It appears that nonsigning children do not have full control of abstract deixis in gesture until age four or five (McNeill, personal communication) and there is evidence that nonsigning adults with right hemisphere damage maintain the use of space for gestures of abstract deixis, while iconic gestures are impaired (McNeill and Pedelty 1995). More studies of the use of abstract deixis in gesture by nonsigners are required in order to be able to fully evaluate whether the psycholinguistic studies of verb agreement in ASL reveal something about the nature of ASL or about the nature of abstract referential pointing gestures.

I believe that a large part of the conflict between the gestural and grammatical accounts results from a misunderstanding of what gesture means in relationship to language and where we are allowed to find gesture. I present McNeill's (1992) notion of gesture as a modality-free notion and show how speakers not only use speech and manual gesture at the same time, but also speech and spoken gesture at the same time. For speakers, gestural and linguistic codes can be combined in the same vocal–auditory channel. However, because they do take place in the same channel, there are restrictions on how that combination is executed. I argue that signers can also combine a gestural and linguistic code in the same manual–visual channel, and that the restrictions on how pointing is carried out need not mean that the location features are morphemic, but rather that there are restrictions on how the two codes combine in one channel.

7.3   The morpheme vs. gesture question

7.3.1   *What is a morpheme?*

So far I have laid out the controversy as follows: the location features are seen as being either morphemic or gestural. Clearly, the direction of the pointing of an agreement verb is meaningful, and forms used in language which are meaningful – and which must combine with other forms in order to be artic-ulated – are traditionally assigned to the level of morphological description. Yet, some meaningful behaviors involved in language use, such as gesture and intonation, have not traditionally been seen as belonging to a morphological level of description; rather, they are traditionally seen as paralinguistic: parasitic on linguistic forms, but not themselves linguistic. So to say that the location features of agreement verbs are nonmorphological implies that they are pushed out of the linguistic, to the paralinguistic level of description.

It is not necessarily the case that meaningful forms that are not morpho-logical must therefore be nonlinguistic. There are approaches which allow for meaningful aspects of form which are not morphological, but are still linguistic. Woodbury (1987), working in the lexical phonology framework (see Kiparsky 1982) challenges traditional duality of patterning assumptions and the use of abstract morphological 'place holders'[5] by proposing an account of meaning-ful postlexical phonological processes. Researchers in intonational phonology view intonation as having categorical linguistic structure (for an overview, see Ladd 1996) apart from any claims about whether the relationship between that structure and the meanings it imparts is a morphological one. These approaches share the view that meaningful phonology need not be morphological, but what

---

[5] This refers to the use of abstract features to stand in for the "meaning" of phonological processes that occur at a later stage of a derivation or a different module of a grammar. These abstract features then have a chance to be manipulated in the syntax without violating the assumption that all phonology takes place separately from levels dealing with meaning.

makes it linguistic as opposed to paralinguistic is the categorical structure of its form. Admittedly, such a statement is a bit of an oversimplification. The matter of deciding which aspects of intonation are linguistic and which are paralinguistic has not been resolved; it is not clear what kind of linguistic status should be ascribed to gradient, final postlexical processes of phonetic implementation, but these approaches generally imply that what makes a form–meaning pairing a linguistic one is the categorical nature of the units out of which the form is built.

I have stated that some linguists have found ways to see some aspects of meaningful phonology as being linguistic without them necessarily being morphological. Morphology means different things to different linguists. It has been placed within the phonology, the lexicon, the syntax, or on its own level by different theorists. In this chapter, I appeal to a general, structuralist definition of the morpheme: a minimal linguistic unit of form–meaning pairing. The question I aim to address – i.e. whether the location features of agreement verbs are morphemes or gestures – hinges on the term "linguistic" in the definition above. I have already mentioned one possible criterion for considering something linguistic or not: that of categorical (as opposed to gradient) form. The categorical–gradient distinction is important in this discussion, but it is not the end of the story.

"Conventionality" is an important concept in the determination of whether something is linguistic or not. The idealized linguistic symbol is conventionalized. This is related to the notion that the idealized linguistic symbol is arbitrary. When the form of a sign does not have any necessary connection with its meaning, if it does not have an iconic or indexical relationship to its referent, the only way it could be connected to its meaning by language users is by some sort of agreement, implicit or explicit, that such a form–meaning pairing exists: convention steps in as an explanation for the successful use of arbitrary signs. However, a sign need not be arbitrary to be conventional. The onomatopoeia *woof* is iconic of a dog's bark, but it is also conventional. The reason that it is conventional is not because there would otherwise be no way to transmit its meaning, but rather because this is the form that English speakers regularly use as a symbol of a dog's bark. It is this regularity of use, this consistent pairing of form and meaning over many instances of use that creates a conventional sign. In Russian *gaf* is the conventional sign for a dog's bark. This form is just as iconic as *woof*.[6] An English speaker could use the Russian form to represent a dog's bark, but to do so he or she would need a lot of contextual support. If he were to utter *gaf* in the acoustic manner of a dog's bark, while panting like a dog and pointing to a picture of a dog, it would likely be understood that *gaf* was meant to signify a dog's bark. *Gaf* is not a form regularly paired in English

---

[6] I appeal to a rather intuitive notion of iconicity here; for a thorough discussion of iconicity, see Taub 1998.

with such a meaning, and although it can still have the meaning of a dog's bark in a specific context of speaking described above, it does not have that meaning by virtue of convention. A conventional sign does not need as much contextual support to have a meaning.

Part of what is needed for a form to be conventionalized is the ability of that form to be used to mean something stable over a wide range of specific speaking events. In Langacker's (1987) terms, it becomes "entrenched" and "decontextualized." It becomes entrenched as it is consistently paired with a meaning and, because of that, it comes to be decontextualized, i.e. it comes to have a meaning and a form that abstract away from the details of any one occasion of its use. It is important to note that a symbolic unit can be more or less entrenched, and more or less decontextualized, but there is no criterial border that separates the conventionalized from the unconventionalized.

Freyd (1983) suggests that linguistic forms are categorical, as opposed to gradient, as a result of their being conventional. Due to "shareability constraints," people stabilize forms as a group, creating categories in order to minimize information loss. A form structured around stable categories can be pronounced in many ways and in many contexts but still retain a particular meaning. A change in form that does not put that form in a different category does not result in a change in meaning. But if a form is gradient, a change in that form leads to a concomitant change of meaning, and the nature of that change is different in different contexts.

A form can be said to be "conventionalized" if it has a stable form and meaning, which have come into being through consistent pairing and regular use. A morpheme is a conventionalized form–meaning pairing and, because of its conventionality, has categorical structure. Also, because of its conventionality, it is decontextualized so it is listable independent from the speech event. It is worth noting here that while the ideal linguistic unit is a conventionalized unit, it is not necessarily the case that every conventionalized unit is linguistic. Emblematic gestures (see Section 7.3.2.1) like 'thumbs up' are quite conventionalized, but we would not want to say they are English words because their form is so vastly different from the majority of English words. Given that a conventionalized unit is not necessarily a linguistic unit, is a linguistic unit necessarily conventionalized? A linguistic unit necessarily involves conventions. The question remains as to what the nature of those conventions must be. I explore this question further in Section 7.5 when I discuss the issues that motivate the criteria for deciding how to draw the line between morpheme and gesture.

7.3.2   What is a gesture?

The controversy over whether agreement is gestural or morphemic depends heavily on what one's notion of gesture is. What does gesture even mean in

a visual–manual language? It cannot mean 'things you do with your hands,' because then all of sign language would be gesture, and we obviously do not want to equate signed language with the hand movements that people make while speaking a spoken language. To answer this question, we need a modality-free notion of gesture.

7.3.2.1 *Emblems.*   I adopt here McNeill's (1992; 1997) notion of gesture. The most common, colloquial use of the term "gesture" refers to emblems. Emblems are gestures such as the "OK" sign, the "goodbye" wave, and the "shushing" gesture. These gestures have fully specified forms: there is a correct way to produce them. These gestures are picked from a set of defined forms, and in that way they are similar to lexical items. They are listable and specified, like words, and can be used in the presence or absence of speech. They are conventionalized.

   I mention these emblem gestures in order to stress that my use of the term "gesture" does not refer to them, and I would like to lay them aside. The notion of gesture that I refer to uses the term to describe speech-synchronized gestures that are created online during speaking. These gestures are not plucked from a gesture lexicon and pasted into the conversation; they rather arise in concert with the speech, and their form reflects the processing stage that gave rise to the speech. The gestures are not conventionalized and are produced with speech.

7.3.2.2 *Speech synchronized gestures.*   Figure 7.1 shows a normal example of speech-synchronized gesture. These pictures are video stills from a research project in which subjects are videotaped telling the story of a Sylvester and Tweety cartoon that they have just seen. If one looks at these gestures without reading the description of the event being narrated (below), there is no way to know what these gestures mean. Unlike the case of emblems or signs, a speech-synchronized gesture does not express a conventionalized form–meaning pairing. The conventionalized pairings occur in the accompanying speech and reliably communicate the meaning to an interlocutor who knows the conventions. The gestures on their own cannot reliably communicate meaning because they are not conventionalized. Their forms are created in the moment and directly reflect the image around which the speaker is building his (or her) speech. However, because those forms do reflect the image around which the speech is built, knowing what that imagery is (either through hearing the speech at the same time, or through knowing the content of the narration) will render the gestures meaningful. In the scene that the speaker is describing in Figure 7.1, Sylvester makes an attempt to capture Tweety, who is in a cage in a window high up on a building. He makes a seesaw with a block and a board, stands on one side of the seesaw, and throws an anvil on the other side, propelling himself upward. He grabs Tweety at the window and falls back down onto his

Figure 7.1   Video stills of speaker telling the story of a cartoon he has just watched

side of the seesaw, propelling the anvil upward. The anvil then comes down on his head, flattening it, and Tweety escapes.

The gestures pictured in Figure 7.1 should now seem quite transparent in meaning. They are meaningful not because they are conventionalized, but because you know the imagery they are based on, and so will see that imagery in the gestures. The representational significance of the gestures (a–l) pictured in Figure 7.1 is given in (1).

(1)   a. introduction of the seesaw
           b. introduction of the weight
       c–d. Sylvester throws the weight
           e. Sylvester goes up

  f. and comes down
g–h. grabs Tweety
 i–j. Sylvester comes down, sending the weight up
k–l. weight comes down and smashes Sylvester

These gestures do not reproduce exactly the events of the cartoon, in the same order. Iconic speech-synchronized gestures are not remembered images recreated on the hands. They are representations "formed within and conditioned by the linguistic system" (Duncan 1998). The speaker packages the incident into the linguistic structures given by his language and the gesture emerges in sync with and congruent with that packaging. So what he does with his gestures is somewhat constrained by the way he has packaged the event for his speech. Examples (e) and (f) represent Sylvester's overall trajectory of being propelled upward then falling downward. Then (g–h) represent Sylvester's grabbing Tweety at the window, something that happened before the up–down trajectory was complete. Examples (i–j) begins at the downward part of the overall trajectory, but this time the focus is on the relative trajectories of the Sylvester and the weight. The speaker focuses on the relative trajectories in order to set up the punchline: Sylvester getting smashed by the weight.

  The speaker did not take these gestures from a list of pre-existing conventionalized gestures. He created them on the spot, in the context of his narrative. They reflect imagery that is important in his discourse at the moment. Also, they do not simply depict the scene. They depict the scene in chunks that his linguistic packaging of the scene has created.

  The gestures pictured above are all "iconics" (McNeill 1992). They depict concrete aspects of imagery with forms that look like the images they represent. There are other classes of speech-synchronized gesture that do not work on exactly the same principle. The most important class to mention for purposes of this chapter is that of "abstract deixis" (McNeill *et al.* 1993). In gestures of abstract deixis, speakers establish loci for elements of the discourse by pointing to the empty space in front of them. Like signers, they point to the same spatial locus when talking about the entity that they have established there, but unlike signers, they are not required to do so consistently. This is due to the fact that the speech is carrying most of the communicative content, leaving more room for referential errors in the gesture.[7]

  This speech-synchronized gesturing is very robust around the world, and it is used by speakers of all languages. This does not necessarily mean that we should expect signers to do it as well, but it does give us a good motivation to look for it. If signers were to have sign-synchronized gesture, they would have to simultaneously gesture and sign.

---

[7] However, there is evidence that interlocutors do pick up on information communicated in gesture that is not present in speech (McNeill *et al.* 1994). It is probable that inconsistency in maintaining loci in gesture leads to comprehension difficulties.

WHY   **WAKE-UP**   **EARLY**

gesture   **ASK-HIM**   **gesture**

Figure 7.2   Illustration of (2)

7.3.2.3 *Less controversial analyses of gesture in sign.*

- Code suspension: Researchers looking at gesture in sign have considered certain manual actions that are interspersed with the sign stream as gestures (Marschark 1994; Emmorey 1999). For example, a thumb-point or a shoulder-shrug may be considered a gesture. These gestures do not, however, happen simultaneously with signs. They pattern as suspensions of the linguistic code: SIGN SIGN GESTURE SIGN, as in (2), which is illustrated in Figure 7.2.

> (2) WHY WAKE-UP EARLY *gesture* ASK *gesture*
>   'Why does he wake up early? Ask him.'

While gestures that interrupt the sign stream are an interesting area for further research, they do not speak to the issue of simultaneously articulated sign and gesture.[8]

- Gesture on different articulators: Some nonmanual actions have been considered candidates for gesture, namely affect-displaying face and body posture

---

[8] I am neutral on whether these emblem-type signs should be considered proper lexical signs or gestures. They have different distributional properties from lexical signs (Emmorey 1999), but it may be simply that we tend not to view forms with purely pragmatic meaning as fully linguistic, for example English *mmm-hmmm*.

LOOK-AROUND   **SMILE**

Figure 7.3   Illustration of (3)

(Emmorey 1999) and "constructed actions" (Metzger 1995). These gestures are articulated simultaneously with signs, but on separate articulators. In (3) the signer signs LOOK-AROUND while himself looking around, and SMILE while smiling (see illustration in Figure 7.3).

(3) ____looking __smiling
     LOOK-AROUND SMILE
     'He looked around and then smiled.'

This kind of example shows the co-ordination of a gesture-producing articulator with a sign-producing articulator, and is similar to the examples of hearing people producing speech on the vocal articulator while producing gesture on the manual articulators. However, the question of whether pointing verbs can contain gesture within them is a question of whether a signer can gesture and sign simultaneously on the same articulator.

7.3.2.4 *Gesture as a semiotic notion.*   The above discussion of simultaneously produced gesture and speech does not seem to reveal anything about the simultaneous production of ASL and gesture on the same articulator. The speech is produced with the vocal articulators while the gesture is produced with the manual articulators, so there can be no confusion as to which is which. However, McNeill's (1992; 1997) notion of gesture is not modality bound. It is a semiotic, not a physical notion. The semiotic notion of gesture is:

• That which expresses the imagistic side of thought during speaking through forms directly created to conform to that imagery. The imagery can be concrete or abstract.

McNeill (1992) argues that in language use two kinds of cognition are in play: imagistic thinking (Arnheim 1969; see also the "spatial-motoric thinking" of Kita 2000) and analytic thinking. In imagistic thinking, concepts are represented as "global-synthetic" images: "global" in that the parts of the image have meaning only insofar as they form part of a meaningful whole, and "synthetic" in that the parts of the image are synthesized into a single, unitary

image. In analytic thinking, multiple separately identifiable components are assembled and put into linear and hierarchical order (McNeill 1992:245). Gestures are viewed as a manifestation of imagistic thinking and speech as a manifestation of analytic thinking. The structure of the forms produced in gesture and in speech reflect the nature of the different thought processes: gestures are global-synthetic and speech is linear, segmented, and hierarchically organized.

Theories of language aligned with the "cognitive linguistics" perspective propose that the structure of all aspects of language is motivated by imagery. The structure of both form and meaning in language is determined by abstract schematic imagery, from word structure and the lexicon, to sentence structure (Langacker 1987), to reference and discourse structure (Fauconnier 1994). It may be argued, from that perspective, that both gesture and speech are manifestations of imagistic thinking. However, gesture is still different in that it manifests imagistic thinking directly. The images themselves are given physical realization in the immediate environment. The grabbing action of Sylvester becomes the grabbing action of a hand; the smashing of an anvil on Sylvester's head becomes the smashing of one hand upon the other. In speech, the representation of imagery is mediated by access to an inventory of symbolic units. Those units are symbols by virtue of convention and need not have forms that directly realize imagery. One may construe the lexical meaning of *hypotenuse* with reference to the image of a right triangle with its longest side foregrounded (Langacker 1991), but there is nothing about the physical form of the utterance [haɪpatɪnjuːs] which itself manifests that imagery.

• The forms created are unconventionalized.

Upon first looking at Figure 7.1, before reading the description of the scene being described, a reader of this chapter (most likely) has no access to the meaning of any of the gestures depicted. Figure 7.1e represents Sylvester being propelled upward. It only has that meaning in this particular speaking event; there is no convention of that particular gesture being paired with such a meaning. One might argue that an utterance of the word *up* likewise could only mean "Sylvester was propelled upward" in a specific context. It is true that words are not completely specified for meaning; some aspects of the meaning of *up* in any one instance of its usage are supplied by the context of the speech event. It could refer to a high static location or a movement, different trajectory shapes, different velocities of movement, etc. The conventionalized meaning of *up* abstracts away from all such particulars, leaving a rather abstract schema for 'up' as its meaning. One might look at the gesture in (1e) and without knowing that it represents Sylvester being propelled upward, recognize it as some sort of depiction of the abstract schema of 'upness.' However, the word *up*, as a conventional sign for 'upness,' may be underspecified with respect to particulars of its meaning in context, but it is completely specified with respect to its

form. In contrast, there is no convention that determines how the gesture for 'upness' should be pronounced. It may, as a gesture, come out as the raising of a fist, the lift of a shoulder, or the raising of the chin. All such gestures do share one feature of form; they move upward, and one might be tempted to say that that aspect of upward movement alone is the conventionalized form for representing upward movement. Such a form–meaning pairing is tautological: 'up' means 'up,' and need not appeal to convention for its existence. Additionally, the spoken phrase "He flew up" can be uttered in sync with a gesture which moves downward without contradiction. For example, the speaker could make a gesture for Sylvester's arms flying down to his sides due to the velocity of his own upward movement. The only motivation for the form of a speech-synchronized gesture is the imagery in the speaker's thought at the moment of speaking.

That being said, there are some conventions involved in the production of gesture. There may be cultural conventions that determine the amount of gesturing used or that prevent some taboo actions (such as pointing directly at the addressee) from occurring. There are also cultural conventions that determine what kind of imagery we access for abstract concepts. Webb (1996) has found that there are recurring form–meaning pairings in gesture. For example, an "F" handshape (the thumb and index fingers pinched together with the other fingers spread) or an "O" handshape (all the fingers pinched together) is regularly used to represent "preciseness" in the discourses she has analyzed. According to McNeill (personal communication), it is not the conventionality of the form–meaning pairing that gives rise to such regularity, but the conventionality of the imagery in the metaphors we use to understand abstract concepts (in the sense of Lakoff and Johnson 1980; Lakoff 1987). What is conventional is that we conceive of preciseness as something small and to be gingerly handled with the fingertips. The handshape used to represent this imagery then comes to look alike across different people who share that imagery. The disagreement here is not one of whether there are conventions involved in the use of gestures. It is rather one of where the site of conventionalization lies. Is it the forms themselves that are conventionalized, as Webb claims, or the conceptual metaphors that give rise to those forms, as McNeill claims? The issue of "site of conventionalization" is also important for the question of whether the pointing in agreement verbs in ASL is linguistic or gestural. I give more attention to this issue in Section 7.5.2.

In any case, what is important here is that the gestures are not formed out of discrete, conventionalized components in the same way that spoken utterances are. And even if there are conventionalized aspects to gesturing, they are far less conventionalized than the elements of speech.

• The form of the gesture patterns meaning onto form in a gradient, as opposed to a categorical way.

In gesture, differences in form correspond to differences in meaning in a continuous fashion. In speech, many different pronunciations of a word are linked to the same conventional meaning.

To sum up, gesture is:

- That which expresses the imagistic side of thought during speaking through forms directly created to conform to that imagery. The imagery can be concrete or abstract.
- The forms created are unconventionalized.
- The form of the gesture patterns meaning onto form in a gradient, as opposed to a categorical way.

With this notion of gesture, many actions produced with the vocal articulators can be considered gesture.

7.4   Spoken gesture[9]

In this section I discuss something I call "spoken gesture." This term covers things that people do with their voices while speaking which exhibit the properties of gesture described above. Example (4) is an example of the meaningful manipulation of vowel length.

(4)   It was a looooong time.

In (4), the word *long* is clearly a linguistic, listable unit, as are the phonemes that compose it. The lengthening of the vowel expresses the imagery of temporal extension through actual temporal extension. The lengthening of the vowel is not a phonemic feature of the word, nor a result of phonotactic considerations. It is not the result of choosing a feature [+long] from the finite set of phonetic features. It is an expression of imagery through a directly created form that happens to be simultaneously articulated with the prescribed form of the vowel.

In (5), the acoustic parameter of fundamental frequency is manipulated.

(5)   The bird flew up [high pitch] and down [low pitch].

The words *up* and *down* are linguistic, listable units, as are the phonemes that compose them. The high pitch on *up* expresses the imagery of highness through

---

[9] All of my examples of spoken gesture are akin to the iconics class of gesture. My argument, insofar as it addresses gesture in agreement verbs, would be better served by examples of spoken gesture that are akin to abstract deixis, but there can be no correlate of abstract deixis in speech because there is simply no way to point with sound waves. One can, of course, refer with sound and words, but the speech channel cannot support true pointing. A reviewer suggested that pointing can be accomplished in speech by using pitch, vowel length, or amplitude to index the distance of referents. I have not quite resolved whether to consider such use of acoustic parameters to be pointing, but for now I will say that the defining characteristic of pointing is that it shows you where to direct your attention to in order to ascertain the referent of the pointing. An index of distance alone gives you some idea of where to direct your attention, but a much less precise one.

a metaphor that associates high vocal frequency with high spatial location. The low pitch on *down* expresses the imagery of lowness through the flip side of that metaphor. The tones are not phonemic features of the words. The tones express imagery through a directly created form which is simultaneously articulated with fixed lexical items.

In (6), repetition is exploited for effect.

(6)   Work, work, work, rest.

The words in this example are linguistic, listable units. The construction in which they occur cannot be given a syntactic–semantic analysis. The quantity of words reflects the imagery of the respective quantity of work and rest. The ordering of the words reflects the ordering of the actions. A lot of work, followed by a little rest. The words are chosen from the lexicon. The construction in which they occur is created in the moment.

The examples of spoken gesture show that speakers can express the linguistic and the gestural simultaneously, on the same articulators, in the same modality.

Liddell does not propose that agreement verbs are gestures. He rather proposes that they are linguistic units (features of handshape, orientation, etc.) simultaneously articulated with pointing gestures (the location features, or "where the verb points"). Spoken linguistic units can be articulated simultaneously with spoken gestures. That signs could be articulated simultaneously with manual gestures is at least a possibility.

7.5   The criteria

I have discussed differences between the morphemic (or linguistic) and the gestural and introduced the idea of spoken gesture. In sections 7.5.1–7.5.3 I discuss the problematic issues that arise when trying to draw a clear distinction between what is linguistic and what is gestural when both types of code are expressed in the same channel, on the same articulators. The issues are:
• "degree of conventionalization" of a form;
• "site of conventionalization" of a convention; and
• "restriction on combination" of a gesture with a linguistic form.
They are presented as the dimensions along which criteria for deciding where to draw the line between morpheme and gesture can be established.

7.5.1   *The determination of conventionalization is a problem*

Put simply, the form of a gesture is unconventionalized, while the form of a word or morpheme is fully conventionalized. However, these examples of spoken gesture present some problems in the determination of their level of conventionalization. The vowel extension in (4) can also be applied to other

words. It can be applied to other adjectives as in *It was biiiiiiig* or to modifiers as in *It was veeeeeeeery difficult*. In all cases the lengthening seems to convey the meaning of "moreness." Are we dealing with a morpheme here? Most linguistic theories would say no. Still, there is something conventionalized about the vowel lengthening. It seems to be somewhat regularly and consistently done, but certainly not as consistently as the form *more* is paired with the meaning of "moreness." Also, unlike the case with conventionalized lexical items or morphemes, we can take advantage of the gradient nature of the patterning in order to make fine distinctions in meaning. We can vary the meaning by correspondingly varying the form in a nondiscrete way. Consider *He had a biiiig house, but his parents had a biiiiiiiiig house*.

The tone sequence in (5) may also be seen as conventionalized within the realm of intonational phonology. The particular sequence of pitch accents and boundary tones is also found in other common collocations such as *back and forth* and *day and night*. However, if gradient patterning is taken advantage of and the intonation is extended to *The bird flew uuuuuup* [with exaggerated rising contour] *and dooooown* [with exaggerated falling contour], then the phrase does not follow a conventionalized pattern. The phrase *day and night* said with the same exaggerated intonation would be odd (although for me, saying it that way evokes the image of a rising and setting sun). Through the manipulation of pitch, intensity, and timing parameters we can continue to map more meaning onto the form by adding force dynamics, speed, and particular flight contours to the articulation, all of which would certainly be outside the realm of intonational phonology. How much of that manipulation can be considered conventionalized?

One could also see conventionalized aspects in example (6). There seems to be some sort of convention to three repetitions, as in *Work work work, that's all I ever do* or *Bills bills bills, can't I get any other kind of mail?* However, as with the examples above, the ordered sequence of words in (9) can also be made more imagistic by taking advantage of gradient patterning. Consider *Work, work, work, work, work* [said very rapidly] – *rest* [said very slowly]. Again, speed, intensity, force dynamics, and other sequencing possibilities can be mapped directly onto the form of the utterance. It is difficult to determine where the conventional leaves off and the idiosyncratic begins.

The purpose of this section has been to show that it is no trivial matter to decide whether something is conventionalized and how conventionalized it is when we leave the extremes of the scale. It is probably reasonable to have degree of conventionalization as a criterion for considering something gestural or linguistic, with completely conventionalized on the linguistic side and completely unconventionalized on the gestural side. Unfortunately, most of the cases we have difficulty with lie in between these extremes.

Degree of Conventionalization: The determination of the degree to which something is conventionalized is certainly useful in deciding whether something is gestural or morphemic. However, it is no trivial matter to make this determination, and the researcher must be wary of depending on this criterion entirely.

7.5.2   *The determination of the site of conventionalization is important*

The controversy over the gesture proposal for agreement verbs is due in large part to the fact that in deciding whether the pointing is morphemic or gestural, different criteria are being used. For Liddell, the form of the locus position being pointed to by the agreement verb is not conventionalized, so it is gesture. The objections stress that the way that pointing in general is carried out is restricted in a conventionalized way, so the pointing is morphemic. Both points of view take degree of conventionalization as a criterion of decision, but they differ on the site of conventionalization which is important: conventionalization of the form of the locus position itself vs. conventionalization of the way in which pointing is carried out in general.

It is clear that there are conventions that govern the use of pointing. If the criterion for considering something gesture is whether or not it is conventionalized, then both viewpoints are correct, but in different ways. It is gesture because its form is nonconventionalized; it is linguistic because the practice of using those forms is conventionalized. What we have here is a disagreement about which stratum of conventionalization is important in considering a phenomenon linguistic.

Site of Conventionalization: When something is said to be conventionalized, or restricted by language-internal considerations, the researcher must be as explicit as possible about the level at which that conventionalization is located. Is it the form that is conventionalized, or a particular aspect of the use of that form? If the researcher decides that conventionalization on a particular level suggests a linguistic (as opposed to cognitive or cultural) interpretation of the phenomenon, he or she should be consistent in considering parallel cases in spoken language as linguistic as well.

7.5.3   *Restrictions on the combination of the gestural and the linguistic*

7.5.3.1 Where the feature is not elsewhere contrastive.   All agree that there are restrictions on the way pointing is carried out that are language particular. There are two ways of looking at this. Either the fact that it is restricted makes all parts of it linguistic, or there is a gestural part and a linguistic part, and the gestural must combine with the linguistic in a restricted way.

This prompts one to ask whether there are any constraints on the way in which the combination of gesture and speech is carried out in general. It appears there are, as in (4) above. In example (4), all three phonemes (/ɪ/ /ɔ/ /ŋ/) are possible

candidates for temporal extension. All are continuants and could be sustained for longer durations. However, the best way to achieve the combination is to extend the vowel.

(7)   a. *lllllllong time
        b. *longngngng time

This example is intended to show that when speech and gesture must combine in the same channel, there are restrictions on the way they may combine. There are linguistic elements that are better suited to carry the gestures than others. In this case, the vowel is the best bearer of this gesture. It is not always the case that the gestural manipulation must combine with the vowel. But there seem to be restrictions determining which kinds of segments are best combined with certain types of gestures.

7.5.3.2 Where the feature is elsewhere contrastive. The spoken gesture examples I have given here have involved manipulation of parameters that are not otherwise involved in categorical contrasts elsewhere in the language. English does not have phonemic vowel length or tone, nor is reduplication a morphological process. Those parameters are, in a sense, free to be gestural. But what happens in a language where the phonetic form used to convey the imagery has phonemic value?

In Mandarin Chinese, the lexical tones for *shang*, 'up' and for *xia*, 'down' are both falling (tone 4). In this case, the lexical specification for tone seems to preclude any gestural use of tone. Indeed, a speaker cannot make the word for 'up' have a rising contour for the expression of imagery as seen in Figure 7.4 (a spectrogram of an English speaker saying *uuuuup and doooown*). The Chinese speaker cannot express the imagery of an upward climb with a sharp pitch rise over the word for 'up' as above, because the lexical word for 'up,' *shang*, must have a falling tone. However, there are other ways for the gesture to emerge. The normal pitch contour for *shang* 'up' and *xia* 'down' is shown in Figure 7.5.

(8)   Ta   pa   shang   you   pa   xia
        He   climb   up   have   climb   down
        'He climbed up and down.'

Both *shang* and *xia* fall sharply in pitch, and the second articulation starts a little lower than the first due to regular sentence declination. Figure 7.6 shows the pitch contour for the gesture enhanced articulation of *shang* and *xia*. This articulation was elicited by asking the speaker to imagine he was telling a story to a child and wanted to put the imagery of upness and downness into his voice.[10]

[10] The speaker also stated that his utterance sounded like something going up and down and that it sounded quite different from simple stress or emphasis on 'up.'
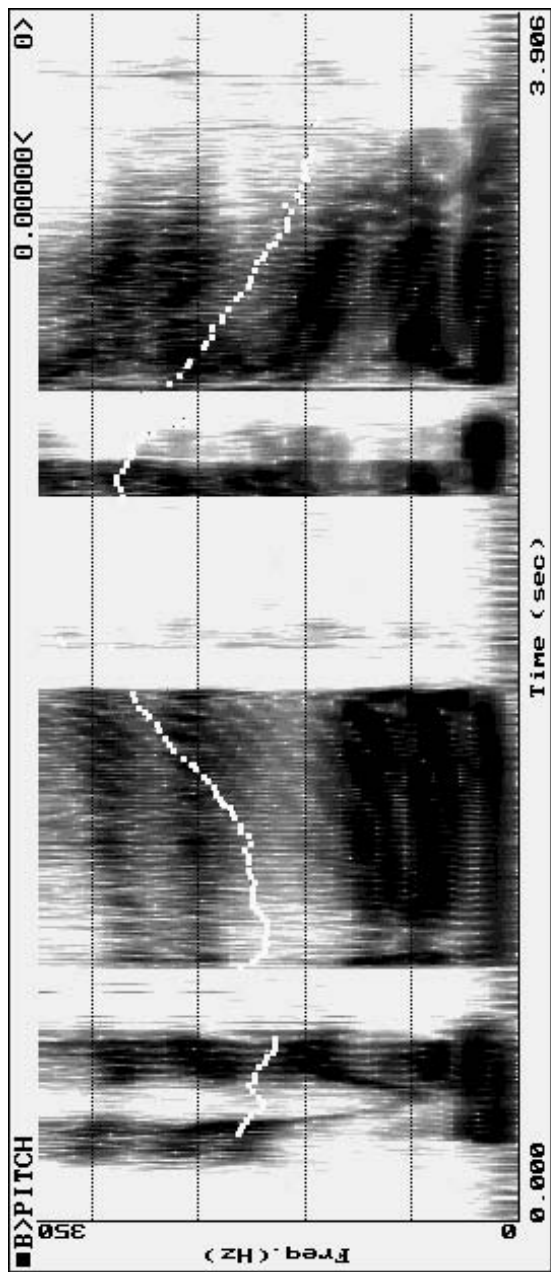
Figure 7.4 Spectrogram of English utterance with gestural intonation

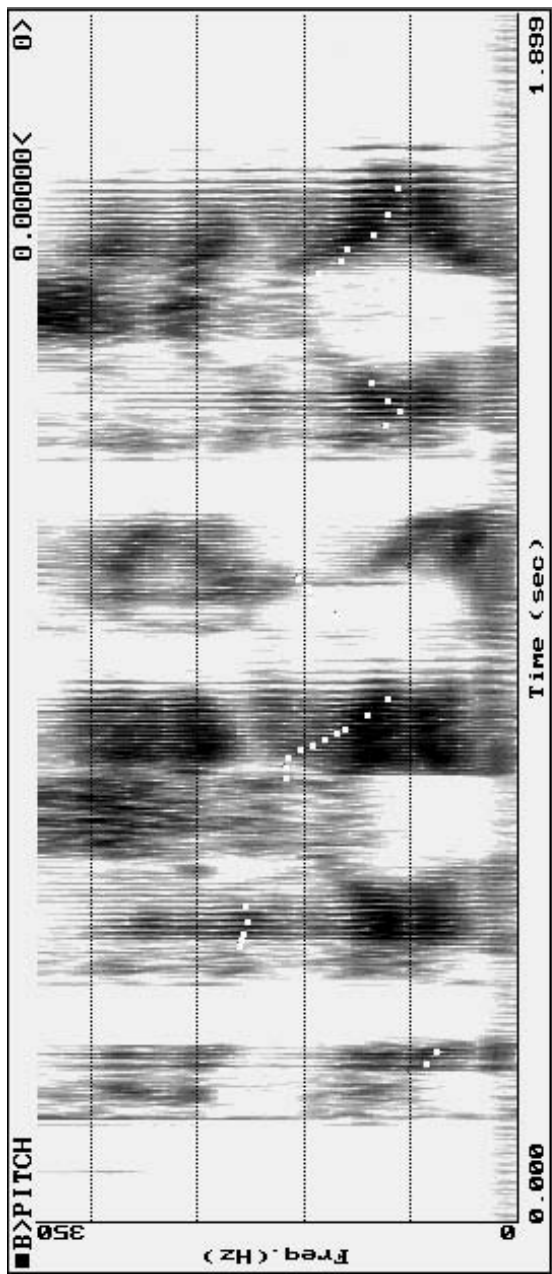He went   uuuuuuuuup   and   doooooown

Figure 7.5　Spectrogram of Chinese utterance with neutral intonation

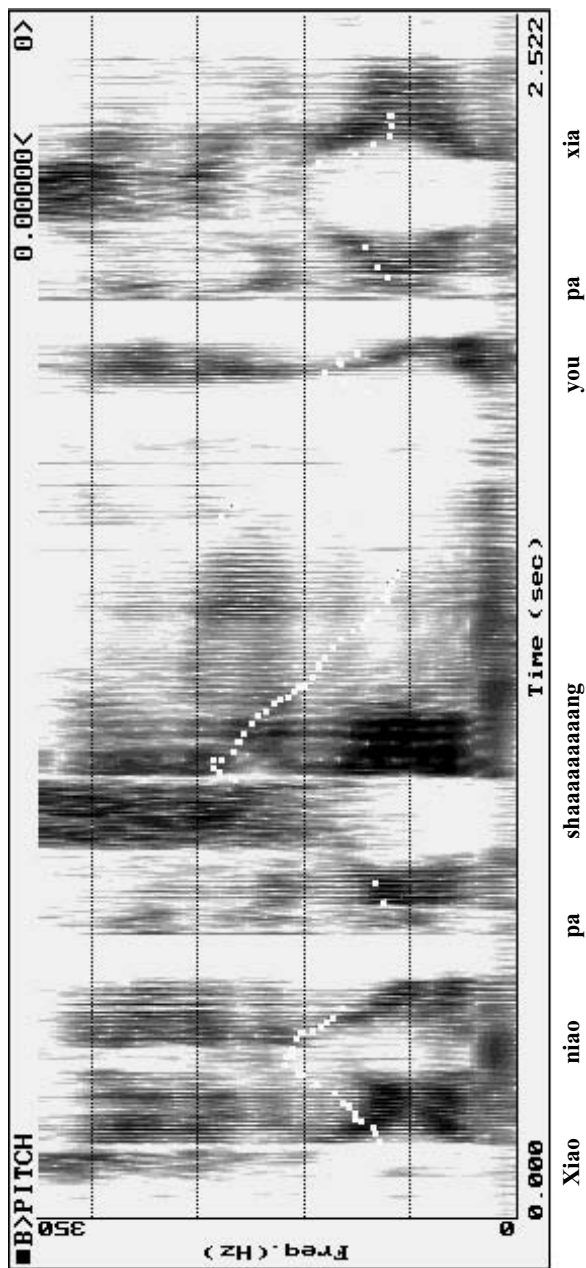Figure 7.6　Spectrogram of Chinese utterance with gestural intonation

(9)   Xiao niao   pa   shang   you   pa   xia
       Bird   climb   up   have   climb   down
       'The bird flew up and down.'

Notice that *shang* does not rise as *uuuup* does in the English example above in Figure 7.4. It cannot rise because of the restriction of lexical tone. However, it is still possible for the speaker to manipulate intonation gesturally. The gesture of 'upness' is achieved through the displacement of the pitch peak of *shang* to be much higher than the peak for *xia* and the extended articulation.

The point of this example is to show that the use of an articulatory parameter for categorial contrast in one area of the language does not render impossible its use in the realm of gesture, and that there may be language-particular restrictions on the way that spoken gesture combines with spoken language.[11]

Not much is known at this point about the kinds of restrictions on the combination of spoken gesture with speech and whether there are any parallels to the restrictions on pointing mentioned as objections to the gesture proposal in Section 7.2. Are there spoken gestures that may only combine with a particular lexical class (first point in Section 7.2)? Expressive reduplication may in some cases be restricted to a certain aspectual class of verbs. Are there spoken gestures which are non-optional (second point)? That depends on whether non-optional intonation patterns that indicate topic and focus, or question and statement, are to be considered gestural. Do different languages handle spoken gesture in different ways (third point)? Chinese and English seem to differ in the way pitch can be used gesturally.

I do not have satisfactory answers to the questions above. At this point I only submit that the dimension of "restriction on combination" is another criterion by which the morpheme vs. gesture question can be evaluated.

Restriction on combination: Restrictions on phenomena can come from the requirements of the grammar, but they can also come from the interplay of two kinds of code upon their integration into one channel. More work needs to be carried out on the nature of the restrictions on gestural forms that result from the requirement that they share the same channel with linguistic forms.

7.6   Conclusions

It is unfounded to reject the idea that agreement is gestural simply because the verbs being produced are linguistic units. People can vocally gesture while

---

[11] There do not seem to be similar restrictions on the combination of manual gesture with speech. Although speech is tightly synchronized with manual gesture – and conveys meaning which is conceptually congruent with speech – the particular forms that the manual gestures take do not appear to be constrained in any way by the specifications on form that the speech must follow. Combining two semiotic codes in one modality may raise issues for language–gesture integration that do not arise for situations where the linguistic and the gestural are carried by separate modalities.

saying spoken lexical verbs. Signers can manually gesture while signing lexical verbs. However, the combination of gesture and speech in one channel puts restrictions on the gesture because important linguistic categorical information, like lexical tone, must be preserved. The three objections given in Section 7.2 above make reference to restrictions on the way in which pointing is carried out, and these restrictions are language particular. The fact that there are language-particular restrictions on the way the pointing is carried out does not in itself constitute a devastating argument against the gesture proposal.

The title of this chapter promises criteria for deciding what is gestural and what is morphemic in ASL linguistics. There is no checklist of necessary and sufficient conditions for membership in either category. There are, however, three continuous dimensions along which researchers can draw a line between gesture and language.[12] I repeat them here:

- The first is "degree of conventionalization." How conventionalized must something be in order to be considered linguistic?
- The second dimension is "site of conventionalization." What kinds of conventions are linguistic conventions?
- The third dimension is "restriction on combination." What kinds of conditions on the combination of semiotic codes are linguistic conditions?

These are not questions I have answers for, but they are the questions that should be addressed in the morpheme vs. gesture controversy in sign language linguistics.

Acknowledgments

References

Aarons, Debra, Benjamin Bahan, Judy Kegl, and Carol Neidle. 1992. Clausal structure and a tier for grammatical marking in American Sign Language. *Nordic Journal of Linguistics* 15:103–142.
Arnheim, Rudolf. 1969. *Visual thinking*. Berkeley, CA: University of California Press.

---

[12] I remain agnostic with respect to whether drawing such a line is ultimately necessary, although I believe that the effort expended in trying to draw that line is very useful for gaining a greater understanding of the nature of communicative behavior.

Bos, Heleen. 1996. Serial verb constructions in the sign language of the Netherlands. Paper presented at the 5th International Conference on Theoretical Issues in Sign Language Research. Montreal, September.

Duncan, Susan. 1998. Evidence from gesture for a conceptual nexus of action and entity. Paper presented at the Annual Conference on Conceptual Structure, Discourse, and Language, Emory University, October.

Emmorey, Karen. 1999. Do signers gesture? In *Gesture, speech, and sign*, eds. Lynn Messing and Ruth Campbell, 133–159. New York: Oxford University Press.

Fauconnier, Gilles. 1994. *Mental spaces: Aspects of meaning construction in natural language*. Cambridge: Cambridge University Press.

Fischer, Susan and Bonnie Gough. 1978. Verbs in American Sign Language. *Sign Language Studies* 18:17–48.

Freyd, Jennifer. 1983. Shareability: The social psychology of epistemology. *Cognitive Science* 7: 191–210.

Kiparsky, Paul. 1982. From cyclic phonology to lexical phonology. *The structure of phonological representations, I*, ed. Harry van der Hulst and Norval Smith, 131–177. Dordrecht: Foris.

Kita, Sotaro. 2000. How representational gestures help speaking. In *Language and gesture*, ed. David McNeill, 162–185. Cambridge: Cambridge University Press.

Klima, Edward and Ursula Bellugi. 1979. *The signs of language*. Cambridge, MA: Harvard University Press.

Ladd, Robert. 1996. *Intonational phonology*. Cambridge: Cambridge University Press.

Lakoff, George and Mark Johnson. 1980. *Metaphors we live by*. Chicago, IL: University of Chicago Press.

Lakoff, George. 1987. *Women, fire and dangerous things: What categories reveal about the mind*. Chicago, IL: University of Chicago Press.

Langacker, Ronald. 1987. *Foundations of Cognitive Grammar, Vol. 1: Theoretical prerequisites*. Stanford, CA: Stanford University Press.

Langacker, Ronald. 1991. Cognitive Grammar. In *Linguistic theory and grammatical description*, ed. Flip Droste and John Joseph, 275–306. Amsterdam: John Benjamins.

Liddell, Scott. 1995. Real, surrogate, and token space: Grammatical consequences in ASL. In *Language, gesture, and space*, eds. Karen Emmorey and Judy Reilly, 19–41. Hillsdale, NJ: Lawrence Erlbaum Associates.

Liddell, Scott. 1996. Spatial representation in discourse: Comparing spoken and signed language. *Lingua* 98:145–167.

Liddell, Scott and Robert Johnson. 1989. American Sign Language: The phonological base. *Sign Language Studies* 64:195–277.

Liddell, Scott. and Melanie Metzger. 1998. Gesture in sign language discourse. *Journal of Pragmatics* 30:657–697.

Lillo-Martin, Diane and Edward Klima. Pointing out differences: ASL pronouns in syntactic theory. In *Theoretical issues in sign language research: Vol. 1*, eds. Susan Fischer and Patricia Siple, 191–210. Chicago, IL: University of Chicago Press.

Marschark, Mark. 1994. Gesture and sign. *Applied Psycholinguistics* 15:209–236.

McNeill, David. 1992. *Hand and mind: What gestures reveal about thought*. Chicago, IL: University of Chicago Press.

McNeill, David. 1997. Growth points cross-linguistically. In *Language and conceptualization*, eds. Jan Nuyts and Eric Pederson, 190–212. Cambridge: Cambridge University Press.

McNeill, David, Justine Cassell, and Elena Levy. 1993. Abstract deixis. *Semiotica* 95: 5–19.

McNeill, David, Justine Cassell, and Karl-Erik McCullough. 1994. Communicative effects of speech-mismatched gestures. *Research on language and social interaction* 27:223–237.

McNeill, David and Laura Pedelty. 1995. Right brain and gesture. In *Language, gesture, and space*, eds. Karen Emmorey and Judy Reilly. Hillsdale, NJ: Lawrence Erlbaum Associates.

Meier, Richard. 1987. Elicited imitation of verb agreement in American Sign Language: Iconically or morphologically determined? *Journal of Memory and Language* 26:362–376.

Metzger, Melanie. 1995. Constructed dialogue and constructed action in American Sign Language. In *Proceedings of the Fourth National Symposium on Sign Language Research and Teaching*, ed. Carol Padden. Silver Spring, MD: National Association of the Deaf.

Newport, Elissa and Richard Meier. 1985. The acquisition of American Sign Language. In *The crosslinguistic study of language acquisition, Vol. 1: The data*, ed. Daniel Slobin. Hillsdale, NJ: Lawrence Erlbaum Associates.

Padden, Carol, 1988. *Interaction of morphology and syntax in American Sign Language*. New York: Garland.

Petitto, Laura. 1987. On the autonomy of language and gesture: Evidence from the acquisition of personal pronouns in American Sign Language. *Cognition* 27:1–52.

Poizner, Howard, Edward Klima, and Ursula Bellugi. 1987. *What the hands reveal about the brain*. Cambridge, MA: MIT Press.

Taub, Sarah. 1998. Language in the body: iconicity and metaphor in American Sign Language. Doctoral dissertation, University of California, Berkeley.

Webb, Rebecca. 1996. Linguistic features of metaphoric gestures. In *Proceedings of the Workshop on the Integration of Gesture in Language and Speech*, ed. Lynn Messing, 79–93. Newark, DE: University of Delaware.

Woodbury, Anthony. 1987. Meaningful phonological processes: A consideration of Central Alaskan Yupik Eskimo prosody. *Language* 63:685–740.