17

The Syllable in Sign Language: Considering the Other Natural Language Modality

Wendy Sandler
Department of English Language and Literature
The University of Haifa
Haifa, Israel

The research program developed by Peter MacNeilage seeks to derive aspects of phonological organization from fundamental physical properties of the speech system, and from there to arrive at reasonable hypotheses about the evolution of speech. Speech is the dominant medium for the transmission of natural human language, and characterizing its organization is clearly very important for our understanding of language as a whole. Speech is not the only medium available to humans, however, and a comprehensive theory of the nature and evolution of language has much to gain by investigating the form of language in the other natural language modality: sign language, the focus of this chapter.

Like spoken languages, sign languages have syllables, the unit that will form the basis for comparison here. As a prosodic unit of organization within the word, sign language syllables bear certain significant similarities to those of spoken language. Such similarities help to shed light on universal properties of linguistic organization, regardless of modality. Yet the form and organization of syllables in the two modalities are quite different, and I will argue that these differences are equally illuminating. The similarities show that spoken and signed languages reflect the same cognitive system in a nontrivial sense. But the differences confirm that certain key aspects of phonological structure must indeed be derived from the physical transmission system, resulting in phonological systems that are in some ways distinct.

The bulk of the chapter is dedicated to a discussion of the syllable in signed languages, pointing out ways in which the unit resembles its spoken language counterpart and also describing how it differs. After a brief introduction to sign language phonology (in Sections 2 and 3), motivation is presented for use of the term, "syllable" in a physical modality that is very different from the oral-aural modality of spoken language. While the notion of 'sentence' or 'word' may be easy to conceive of in a manual-visual language, 'syllable' takes more convincing, which is what Section 4 attempts. The sign language syllable is distinguished from other units such as the word and morpheme, and phonological evidence is presented for the reality of the syllable as a phonological and prosodic element. Differences between the syllable in spoken and signed language are highlighted in Section 4.3. One important difference is the presence of an oscillating mandible as a syllable frame in spoken language (MacNeilage, 1998), and the absence of a comparable frame in sign language. Another is the availability of greater articulatory range to the primary articulator in sign language—the hand(s)—than to what might be considered the spoken language counterpart, the tongue. I will suggest in Section 5 that phonetic differences underlie phonological differences, providing support for MacNeilage and Davis (2000) position that (some of) the phonological form of language is determined by the physical system.

The evolutionary context that motivates this volume requires us to ponder the implications of the descriptions and analyses to be presented. Section 6 provides some remarks on this issue that grow out of the discussion that precedes it. The fact that there are differences in syllable (and other phonological) organization in the two language modalities does not imply that the oral and manual modalities are mutually exclusive. They have too much in common to sustain that view. Nor does it require us to assume that the medium of transmission is essentially extraneous to the structure and organization of language. The latter view is refuted by the fact that some phonological structure clearly derives from the physical properties of the system and is therefore different in each modality. Instead, we need a theory that explains both

commonalities and differences in phonological organization, differences that were chiseled out of the raw material in each modality. And the theory needs a plausible scenario for how this language capacity could have evolved.

Such a theory must not only explain our species' unique endowment for a complex linguistic system; it must also explain our extraordinary capacity to use two different systems. I suggest bimodalism as a starting point for developing a comprehensive theory of the kind described. Specifically, natural languages in the two modalities evolved from complementary aspects of the **same** system, and bimodalism is still apparent in each kind of language if you know where to look.

Two Kinds of Natural Language

Sign languages are natural languages arising spontaneously wherever a group of deaf people has an opportunity to gather and meet regularly (Klima & Bellugi, 1979; Sandler et al., 2004; Senghas et al., 2004). Sign languages are not contrived communication systems, nor is there a single, universal sign language. Instead, there are hundreds of natural sign languages in deaf communities worldwide (Woll et al., 2001). Sign languages are acquired by children in the same stages and time frame as spoken languages (Meier, 1991). Both deaf and hearing children acquire sign language natively if sign language is the primary language in the home. Signed and spoken languages share many significant linguistic properties at all levels of structure (Sandler & Lillo-Martin, 2002, 2006). Certain key areas of the brain are active in the control of spoken and sign language (Emmorey, 2002). And sign languages subserve the same full range of communicative functions as spoken languages, including artistic forms such as poetry (Klima & Bellugi, 1979; Sandler & Lillo-Martin, 2002, 2006; Sutton-Spence & Woll, 1999). A large body of literature on the topic demonstrates that sign languages are full and complex languages with rich expressive capabilities. It is safe to conclude, then, that speech and language are not synonymous. Instead, speech is one primary medium for language, and sign is another.

Sign Languages Have Phonology

The first strictly linguistic investigation of sign language was that of William Stokoe (1960), working on American Sign Language (ASL). That work was seminal because it established a characteristic of sign languages that makes them clearly comparable to spoken languages, a characteristic that is perhaps least expected a priori. That property is duality of patterning (Hockett, 1960). This means that languages have both a meaningless level of structure—in spoken languages, a finite list of sounds which combine in various constrained ways to make words—and a meaningful level of structure—words and sentences. Duality is a fundamental property of language, responsible for its capacity to generate an enormous lexicon of meaningful words from a very small number of meaningless building blocks. Despite their iconic and gestural origins, Stokoe showed that there is also a meaningless level of structure in a sign language, and inaugurated the field of linguistic research on sign language. He showed that signs are not holistic gestures, as they may appear to be at first glance, but rather that they are made up of a small and finite set of meaningless components. Subsequent research showed that there are constraints on the ways in which these components combine to create the words of sign languages (e.g., Battison, 1978; Mandel, 1981), and that the form of a word may change in predictable ways in different morphophonological contexts (e.g., Liddell and Johnson, R., 1986; Sandler, 1987). Together, these discoveries demonstrate that sign languages have phonology. With a meaningless phonological level, sign languages have the building blocks of a potentially large lexicon.

On the basis of minimal pairs formed by substituting a single feature of handshape, location, or movement in a sign, Stokoe proposed these categories as the three basic formational parameters of signs. Stokoe proposed that each handshape, location and movement of the ASL inventory should be compared with a spoken language phoneme. However, he believed that the inter-organization of these elements within a sign is different from the spoken word, that the elements occur simultaneously, not sequentially like consonants and

vowels in spoken words. Other researchers have found evidence of some sequential structure within a sign, although it is more limited than is typical in spoken words (Liddell, 1984; Sandler, 1989). My own work adopts nonlinear theories of phonology and morphology (e.g., Goldsmith, 1976; McCarthy, 1981) to create a model of sign language structure that reveals both simultaneous and sequential properties (Sandler, 1986; 1989; Sandler & Lillo-Martin, 2006). The analysis of the syllable presented in Section 4 assumes this model.

Simultaneous and Sequential Structure in the Sign

In a typical sign, the hand, in a particular configuration, begins in one location and ends in another. For example, JUST-THEN, from Israeli Sign Language (ISL), is illustrated in Figure 17.1. The hand configuration (HC) is

The hand begins at a location (L) a proximal distance above the nondominant hand, and moves (M) to a location in contact with that place of articulation. A partial representation of JUST-THEN is shown in Figure 17.2.

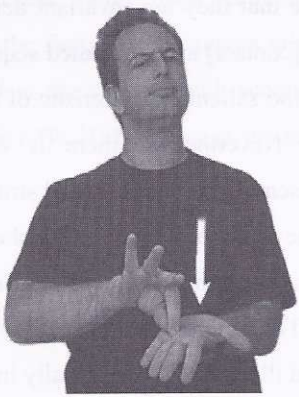


Figure 17.1.
JUST-THEN (ISL)

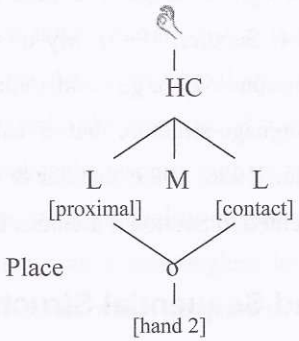
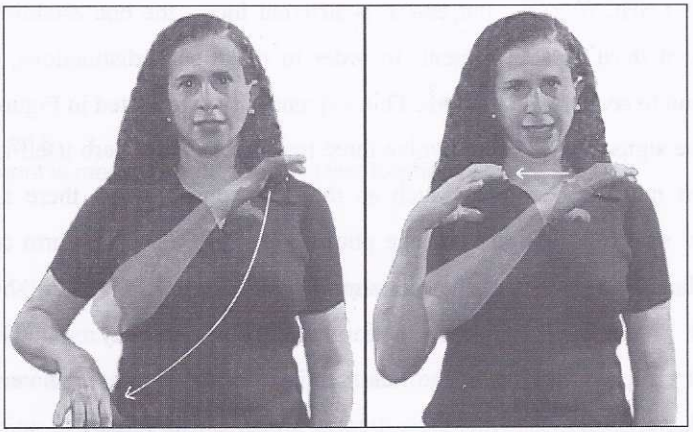


Figure 17.2.
Partial representation of JUST-THEN

The model used for the representation is the Hand Tier model (Sandler, 1986, 1989), motivated by the interaction of sequential and simultaneous elements in sign language phonology and reveals both in its representation. In JUST-THEN both the Hand Configuration and the place of articulation are simultaneous, in the sense that they are invariant across the sign. But the two locations, [proximal] and [contact] are articulated sequentially. While sequential structure is an important and salient characteristic of spoken language, it is less obvious in most signs. Nevertheless, there is compelling evidence that sequentiality is indeed present in the phonological structure of signs.

First, there are some minimal pairs distinguished only by one feature in a sequential position within the sign. Like *chap* and *chat* in English, the signs CHRISTIAN and COMMITTEE are distinguished by the final segment only, as pictured in Figure 17.3 and illustrated schematically in Figure 17.4.

While these minimal pairs, distinguished by sequentially occurring features, are admittedly rare, more evidence for sequentiality is found in the morphophonology of sign languages that have been studied. For example, verb agreement is marked on the first and last locations of a sign. The hand begins at



a. CHRISTIAN b. COMMITTEE

Figure 17.3.
Minimal pair in ASL distinguished by sequentially occurring features.

a location designated as the spatial locus for one referent, typically the subject, and ends at a locus designated as the locus for another referent, typically the object (Liddell, 1984; Padden, 1988). Two pictures of the ISL verb SHOW appear in Figure 17.5. In the first, SHOW agrees with first person subject and second person object, and in the second, with second person subject and first person object. To sign I-SHOW-HER, the sign would begin at the first person

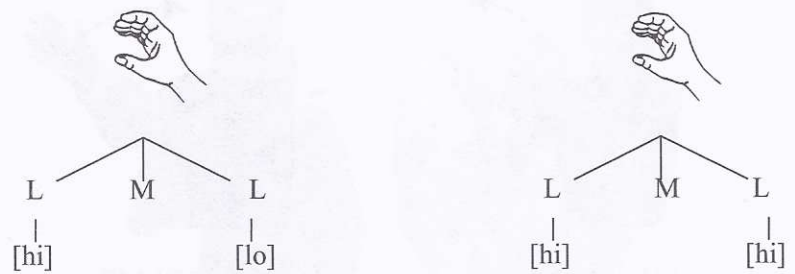


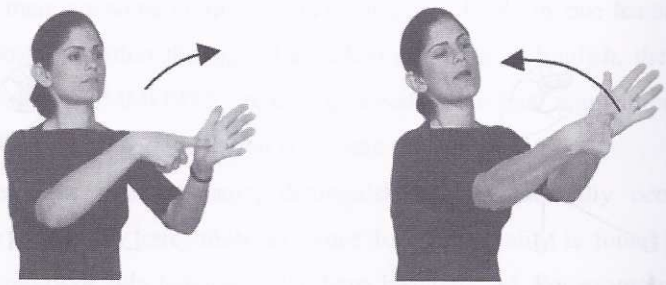
Figure 17.4.
Schematic representation showing features distinguishing CHRISTIAN and COMMITTEE in the final segment of each word.

locus like I-SHOW-YOU, but end at a different locus, the one established for the relevant third person referent. In order to make such distinctions, signers must attend to sequential structure. This sequentiality is reflected in Figure 17.6.

The signs in 17.5 each involve three morphemes: the verb itself and two agreement markers. Examples such as these demonstrate that there is some sequential structure in sign language phonology. Still, the basic form of these signs is the same as that of the monomorphemic sign JUST-THEN, shown in 17.1 and 17.2 above, despite their morphological complexity. All have the canonical structure, LML. In the next section, this structure is shown to be monosyllabic.

The Syllable in Sign Language

Much of the following material summarizes a more detailed treatment in Sandler and Lillo-Martin (2006). In order to demonstrate convincingly that it is useful to adopt the term "syllable" in the description of visually perceived languages, we must show that the unit labeled bears significant similarity to the syllables of spoken languages. This section will demonstrate that syllables of sign language are the anchor to which lower meaningless elements are bound, that they are required to explain morpho-phonological processes, and that they are prosodic in nature. In all these ways, they are like spoken language syllables. I will also



a. I-SHOW-YOU b. YOU-SHOW-ME

Figure 17.5.
Verb agreement

L M L
| |
[agr] [agr]

Figure 17.6.
Agreement is marked on the first and last locations.

describe the differences that also exist between syllables in the two modalities.

The sign language syllable is defined as a single movement, be it path movement of the hand(s) from one location to another, internal movement (such as opening or closing of the hand), or both simultaneously (Brentari, 1998). An example of a sign with path movement only is ISL JUST-THEN, pictured in Figure 17.1.

To argue that sign languages have syllables, it is first necessary to distinguish the syllable from other kinds of structure, such as the morpheme or word. The two sign languages, ASL and ISL, each have many words that are both monomorphemic and monosyllabic, like JUST-THEN shown in Figure 17.1. Monomorphemic but disyllabic words like ISL REVENGE, shown in Figure 17.7 can also be found.

In the verb agreement example in Figure 17.6, a word may consist of several morphemes but still be monosyllabic. And finally, bimorphemic words, like many compounds or words with sequential affixes like ISL SEE-SHARP

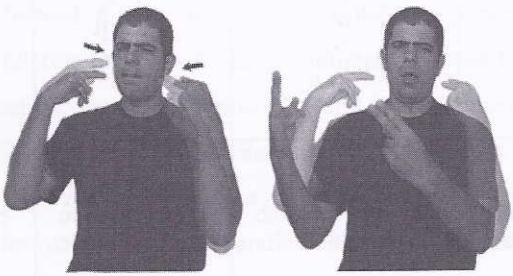


Figure 17.7.
Monomorphemic disyllabic sign: ISL REVENGE

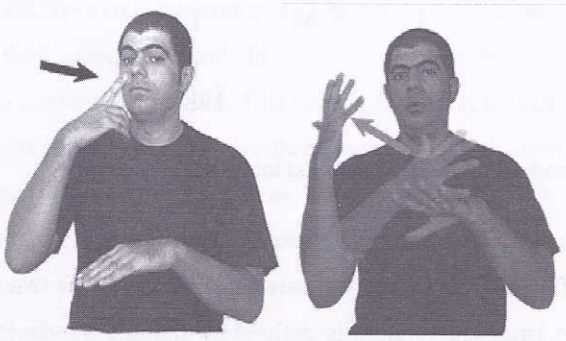
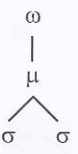


Figure 17.8.
Bimorphemic disyllabic sign: ISL SEE-SHARP ("to discern by seeing").

("discern by seeing") shown in Figure 17.8, may also be disyllabic.

The different relationships between syllables and meaningful units that are found in sign languages are summarized in Table 17.1. By far the most common kinds of words across sign languages are the first and third shown in bold in Table 17.1, i.e., words that are monosyllabic regardless of morphological structure. Forms that are considered to have more than one syllable for the purposes of Table 17.1 are only those that have two different syllables; reduplicated forms are not included. The relation between syllable and word

Table 17.1.
The word, the morpheme, and the syllable are distinguished by their co-occurrence patterns. All the possibilities shown are attested, but those in bold are most common.

			
monomorphemic monosyllabic	monomorphemic disyllabic	**bimorphemic monosyllabic**	bimorphemic disyllabic

reveals a clear modality effect in most spoken languages, especially those with morphological complexity.

In sign language, despite the non-isomorphism between word and syllable, there is an overwhelming tendency for words to be monosyllabic (Coulter, 1982). I refer to this as the monosyllable conspiracy (Sandler, 1999a).

We can see this conspiracy at work where morphologically complex words that either diachronically or underlyingly have more than one syllable reduce to the canonical monosyllable. Lexicalized compounds in sign languages provide an example. In all languages, compounds may become lexicalized where the juxtaposition of the two words is entered as a single word in the lexicon. One indication that a compound has become lexicalized is a meaning that cannot be predicted from the words from which it was originally formed. For example, a *blackboard* does not necessarily have to be black. Another is phonological change, such as fusion. The Hebrew word for soccer is the lexicalized compound *kaduregel* from words meaning "ball" (*kadur*) and "foot" (*regel)*. In the compound, we do not get a geminate (lengthened) [r] sound from the adjacent [r]s in the two words (**kadurregel*); instead, one [r] is deleted.

Returning to the monosyllable conspiracy in sign language, we see an example in one of the lexicalized ASL compounds for the concept FAINT, formed from MIND+DROP, pictured in Figure 17.9. An example of a sign with handshape change and path movement together is DROP, shown in Figure 17.9. MIND and DROP each consists of one syllable in isolation, but in the compound FAINT, the form is not disyllabic as simple concatenation of the two words would predict. Instead, it reduces to a single syllable, represented in 17.10. The autosegmental relation between the hand configurations and the locations under compound reduction, shown in Figures 17.9 and 17.10 is one of the phenomena that motivated the Hand Tier Model (Sandler, 1987, 1989).

Many lexicalized compounds in ASL and ISL reduce to one syllable, and some affixed forms do as well (Sandler, 1999a). We witness a conspiracy toward monosyllabicity in this phenomenon when it is taken together with the overwhelming preponderance of monosyllabic simple words in sign language

a. MIND b. DROP c. FAINT

Figure 17.9.
Hand configuration assimilation in an ASL compound.

lexicons, as well as the tendency for productive morphological processes such as verb agreement to produce monosyllabic words as well. Sign languages seem to prefer monosyllabic words. But in order to justify the existence of the syllable as a phonological and prosodic unit, additional evidence is needed.

Evidence for the Syllable

The first piece of evidence for the syllable as a prosodic unit, then, is the mere fact that signs with one movement (or two simultaneously) are the optimal form in sign languages. As the syllable is not isomorphic with the word (see Table 17.1), the fact that this particular prosodic structure predominates gives us a reason to refer to it in describing the structure of the sign. Several other pieces of evidence for the syllable in American Sign Language have been proposed. Brentari and Poizner (1994) provide evidence that the syllable is a unit of

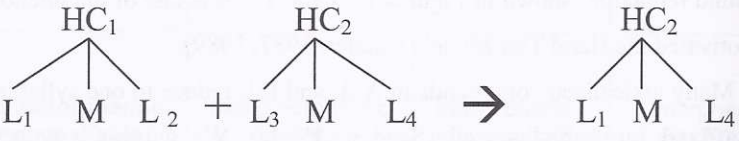


Figure 17.10.
Two syllables reduce to one, producing the canonical, monosyllabic form of a sign (Sandler, 1989, 1999a).

phonological organization by showing that timing of handshape change is different within a syllable than during transitional movements between syllables. The handshape change in a sign like DROP shown in Figure 17.9 is coordinated with, and evenly distributed over, the beginning and ending of the syllable, demarcated by the two locations. However, the handshape change that obligatorily occurs phonetically during the transitional movement between signs is not so coordinated with the last location of one sign and the first location of the next, neither in timing nor in relative distribution of finger movement.

Another reason to believe in syllables is stress assignment in disyllabic signs. Most newly formed compounds and some lexicalized compounds retain the two syllables that are underlyingly present in the two member signs. In such ASL compounds, the second syllable is stressed (Klima & Bellugi, 1979). In contrast, ASL nouns derived through reduplication have stress on the first syllable (Supalla & Newport, 1978).

It is not only stress assignment rules that make reference to the syllable. When ASL verbs are reduplicated under aspectual inflection, the reduplication rule copies only the final syllable (Sandler, 1989). Specifically, if a compound is monosyllabic like FAINT, the whole compound will reduplicate under a temporal aspect inflection such as Habitual (to derive "faint habitually"). But if the compounds have not reduced to a monosyllable and remain disyllabic, like the ASL compound BLOW-TOP (literally, HEAD+EXPLODE-OFF, meaning "explode with anger," only the final syllable undergoes reduplication in the Habitual form. It is clear that these phenomena, (see Table 17.2) are singling out

Table 17.2.
Evidence for the syllable in American Sign Language

1. The optimal form of the sign is a monosyllable (Coulter 1982, Sandler, 1989, 1999a)
2. Handshape change is organized by the syllable unit (Brentari and Poizner, 1994)
3. The final syllable of compounds receives stress (Klima and Bellugi, 1979)
4. The first syllable of reduplicated nominals receives stress (Supalla and Newport, 1978)
5. The final syllable of verbs is reduplicated for temporal aspect inflection (Sandler, 1989)

a prosodic unit, the syllable, not a morphological or lexical unit. It is specifically the rhythmic aspect of the syllable unit at work in each of these constraints and processes, and rhythmicity is prosodic by definition.

Similarities between Spoken and Signed Syllables

Three central characteristics of sign language syllables make them comparable to syllables in spoken language. First, syllables organize lower units of phonological structure. In spoken language, syllables are organized around the nucleus, typically a vowel, and the surrounding consonants usually rise in sonority before the nucleus and fall in sonority after it. Different languages have different constraints on the number of consonants that can occur in the onset and the coda, and on the relative distance in degree of sonority that must exist between adjacent consonants. English clusters that begin with a stop can maximally be followed by one other consonant, which must be a liquid or glide (e. g., *proud, plus,* and *puce,* not **pnack* or **pfack*). In addition, phonological rules may refer to syllables or syllable positions. For example, one of the environments for stop aspiration in English is the onset of stressed syllables.

Now we return to sign language. As we have seen, the timing of handshape change is controlled by the syllable. Although the shape of the hand usually changes in the transitional movement between signs, that change, which is not within a syllable, is uneven in shape and in timing, which leads to the conclusion that the syllable organizes the timing of the units it contains. Using a different model of sign phonology from the on assumed here, Brentari (1998) argues further that all phonological elements that are dynamic have the syllable as their domain.

Second, in neither modality is the syllable unit isomorphic with morphosyntactic structure. It is not the word or the morpheme that is reduplicated in verbal aspect inflection, but the syllable. Similarly, it is the syllable and not the morpheme that receives stress in nominals derived through reduplication.

Finally, syllables in both language modalities are prosodic units. We can see this by their participation in rules and processes that are themselves prosodic in nature, such as reduplication (McCarthy & Prince, 1986) and stress assignment. In fact, it is the prosodic property of "one-movementness" that defines the optimal phonological word in sign language (Sandler, 1999a), and not properties of any nonprosodic unit such as morphemes or lexemes. These observations identify a universal of human language, regardless of modality: a prosodic level of structure that is relevant for linguistic organization and rules, but that cannot be subsumed as part of the morphosyntactic system. Prosodic constituents at higher levels have also been shown to exist in sign languages: the phonological word, the phonological phrase, and the intonational phrase in ISL.

Differences

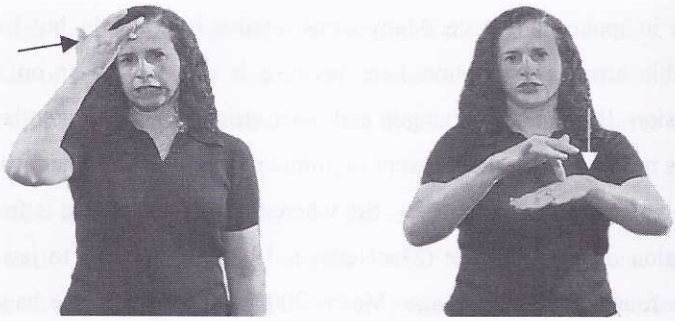
Considering the fundamental lack of similarity in modality of transmission, it is quite striking that the phonological organization of spoken and signed languages should share a prosodic unit at the sublexical level of structure—the syllable. But there are differences as well. Differences in the physical properties of the manual-visual system have reflexes in the organization of the syllable and its role in the phonology.

Because of its many degrees of freedom in the articulation of signs, the primary articulator of sign language, the hand, is sometimes compared with the tongue in spoken language. Many signs involve both hands, but I do not deal with this articulatory option here because it does not bear on the present discussion But unlike the tongue and other articulators of spoken language, the hand is not framed by the inherent rhythmic properties of another articulator that might be compared with the jaw. So, where the spoken syllable is framed by the oscillation of the mandible (MacNeilage, 1998), no parallel to jaw oscillation can be found in sign language (Meier, 2002). In addition, the hand surpasses even the tongue in its articulatory range (Sandler, 1989). First, different combinations of fingers can be selected, for example:

Second, most of these groups can be configured in one of four different positions. Demonstrated here only with the all-five fingers group, positions are respectively open, closed. bent or curved:

Third, the hand can be positioned in any of several different orientations; two examples are:

Finally, the hand can touch or approximate any of a large number of places of articulation on the body. I consider only places of articulation in relation to the body that can be considered system internal and ignore those places of articulation in space. Whether these spatial places are truly linguistic entities is a matter of current controversy (see Sandler and Lillo-Martin, 2006, for discussion). The ASL signs SICK and TOUCH in Figure 17.11 illustrate just two such places. The Hand Tier model (Sandler, 1989; Sandler & Lillo-Martin, 2006) proposes four major body areas—the head (e.g., Figures 17.8—left side, (17.10—left side), the trunk (e.g., Figure 17.3)—the nondominant hand (e.g., Figures 17.1, (17.10—right side) and the nondominant arm—and nine more



a. SICK **b. TOUCH**

Figure 17.11.
Two different places of articulation (ASL)

specific "settings" (such as [hi]—[contralateral], etc.) at each of those major areas. Figures 17.9 above and 17.11 below illustrate two out of the nine possible different settings on the head, ipsilateral in the sign DROP illustrated in Figure 17.9, and central in the sign SICK, illustrated in Figure 17.11.

So even a global comparison between the hand and the tongue is very rough indeed, as the hand has many more degrees of freedom, and it is not grounded within a constricting and oscillating articulator like the jaw.

The phonetics and phonology of the sign language syllable are different from those of its oral counterpart in other ways as well. Unlike spoken syllables in many languages, sign language syllables cannot have clusters of two different locations which might be compared to consonant clusters. Due to the nature of the system, there must be a movement between any two different locations. Similarly, any path movement must by definition traverse the space between two locations, so that it would also be difficult to argue for movement clusters (diphthong-like entities) within a single syllable. Another characteristic of the spoken syllable absent in the sign syllable is an asymmetry between the onset and the rhyme, both in terms of constraints on the constituents (the rhyme is more limited in type and number of segments) and in terms of the role each plays in the phonology (stress assignment cares about the weight of rhymes but not of onsets). Unlike spoken syllables, the syllables of sign language exhibit no onset-rhyme asymmetries; the first and last L do not differ from one another in their articulatory properties or in the role each plays in the system.

In spoken languages, syllables are relevant for the distribution of intonational tunes. Typically, the tunes are aligned with stressed syllables, either within a focused constituent or at a prosodic constituent boundary. While it has been demonstrated that sign languages have intonational systems, conveyed by facial expression, the unit with which intonational tunes are aligned is a larger prosodic constituent, such as the whole phonological or intonational phrase, and not a single syllable within it, stressed or otherwise (Nespor & Sandler, 1999; Sandler, 1999b). Such "tunes" in sign language have been given the label, *superarticulatory arrays* (Sandler, 1999b).

The role of sonority or acoustic resonance in determining the internal organization of the syllable is another important characteristic of the spoken syllable that has no clear analogy in sign language. Spoken syllable onsets rise in relative sonority toward the peak, the syllable nucleus (typically the vowel), and their codas fall in sonority from there, yielding syllables like *plans*, and not like **lpasn*. While several researchers have proposed that sign languages have sonority in the form of relative visual salience (e.g., Brentari, 1990, 1998; Perlmutter, 1992; Sandler, 1993), and even that this relative salience has an effect on the internal structure of the syllable, it is unlikely that useful comparisons can be made regarding a relationship between sonority and syllable organization in the two language systems (see Sandler & Lillo-Martin 2006 for an explanation). The difficulty in finding a parallel stems from a fundamental difference in the architecture of the two transmission systems. In spoken language, the source of energy is the lungs. Relative sonority of the acoustic signal is determined by properties of the vocal tract filter. Sign language has no such distinction between signal source and filter: the signal is perceived directly.

Adding these differences to the other differences in sequential structure outlined above, such as the impossibility of complex onsets, nuclei, or codas, leads to the conclusion that there is no direct analogue to syllable nuclei and margins (vowels and consonants), and that relative sonority is not likely to play a role in sign language syllable organization that is analogous to its role in spoken language. This position, which contrasts in some ways with my own earlier work (Sandler, 1989, 1993), is expanded in Sandler and Lillo-Martin (2006).

Constructing a Lexicon: Less Feature Variegation within the Sign Syllable, but More Phonetic Features in the System

Previously, evidence was presented for sequential structure in sign. However, the segmental structure of sign language is different from spoken language in the following way: most of the features in a monosyllabic sign always characterize all of its segments. It is this broadness in scope of most features that gives the

sign its simultaneous feel. The sign JUST-THEN, pictured in Figure 17.1 and represented schematically in Figure 17.2 illustrates this characteristic.

In the three segments of *fit* [fIt], there is a good deal of variegation in the features and feature values from segment to segment. In addition, few of the features and feature values of any one segment are predictable from the features in the other segments. For example, the rhyme, [It], could easily occur with a different onset, such as [+voiced, +sonorant, +nasal], as in *knit* [nIt]. Or, the onset and nucleus of *fit* could occur with a different coda, such as a voiced lateral sonorant, to produce *fill* [fIl]. The vowel could easily have different features as well, e.g., [+low, -back], to produce *fat* [fæt]. That is, for any feature and feature value in one segment, the features and their values in the other segments are largely unpredictable. And none of the features and values are the same throughout the three segments. While a feature like [voice] or [high] may have the same value throughout a syllable (as in *deal* or *king*, resp.) typically most of the other features will be different. The overall impression is of a sequence of three different segments.

In contrast, in the typical sign, JUST-THEN, almost all the features and their values are the same in the three segments. In all three segments, the index finger is selected and closed (touching the thumb). The palm is oriented downward. The place of articulation is the nondominant hand (h2). Only the features [proximal] in the first segment and [contact] in the last segment differ. While in the English word fit, there are no features that characterize more than two adjacent segments, in the ISL sign JUST-THEN, almost all feature specifications characterize all three segments. This is not an accident associated with this particular sign. Typically there is variation in only one feature in the segments within a sign language syllable. Because so much is the same throughout the sign language syllable, the overall impression is one of simultaneity rather than sequentiality. Some researchers have argued that constraints on production, perception, and short-term memory conspire to create simultaneity of linguistic structure in sign language (e.g., Bellugi & Fischer, S., 1972; Emmorey, 2002).

Signs typically have only one syllable and share most of the same features within that syllable. In principle, this characteristic might limit the potential a sign language has for creating a large number of phonologically distinct and developing a large enough lexicon for adequate communication. Another modality difference may resolve this potential limitation; the number of phonological features available to each system. Comparing phonological models that propose a universal set of features for each modality, sign languages have many more phonological features than spoken languages. Halle (1992) proposes that spoken languages use 18 phonological features to make all the distinctions of their phonological inventories, while Sandler and Lillo-Martin (2006) propose that sign languages require 30, a set almost twice as large as spoken language. Other models of sign language phonology propose larger numbers of features. Other sign language phonologists have motivated different feature inventories, but none of them smaller than 30. Brentari's (1998) carefully detailed model based on American Sign Language proposes 46 features, and van der Kooij's (2002) model of Sign Language of the Netherlands, which strives to minimize redundancy, proposes 39.

An interpretation of these facts is inspired by work by Nettle (1995) comparing ten languages on the basis of two variables, the size of the segment inventory and the length of the word. He found a significant correlation between the two: the smaller the segment inventory, the greater the mean word length. The languages at the two extremes were Nahuatl and !Xu. Nahuatl has an inventory of 23 distinct segments and a mean word length of 8.69 segments, while !Xu has 119 segments and a mean word length of 4.02. Presumably, all 147 segments in !Xu can be distinguished using Halle's 18 distinctive features.

The explanation is simple, and lends itself neatly to the issue at hand. The correlation found by Nettle is compensatory. All natural languages are faced with the same cognitive requirement to furnish a very large lexicon. This can be achieved either by providing a large enough pool of distinctive segments to choose from, or by providing long enough words to enable different combinations of segments in a string. We may extend this line of reasoning to

the somewhat different but comparable issue of syllable internal variegation and feature inventory in signed and spoken languages. Spoken languages have a relatively small number of features but many options for variegation, in this case, for different feature combinations across a syllable (even a syllable with a small number of segments, like *fit*). Sign languages, on the other hand, have a large number of features but very limited variegation across a syllable. According to this reasoning, the limited variegation within a sign syllable is compensated for by the large number of features available for constructing syllables.

The Relation Between the Physical System and Phonology

Many qualitative and quantitative differences in the nature and organization of the syllable in the two natural language modalities have been demonstrated. Differences are attributed to the nature of the physical system of transmission. In spoken language, the syllable frame is provided by jaw oscillation, and content is provided by different configurations of the tongue and lips within the confines of the frame. In sign language, there is no frame to constrain the range or rhythm of the syllable, and the hand articulator has many more degrees of freedom for configuration, movement, and articulation. This added freedom results in a larger number of phonological features in sign than spoken language phonology, a capacity counterbalanced by a limited amount of variegation within a syllable.

These differences between sign and the spoken language syllables provides support for MacNeilage and Davis's research program seeking to derive phonological properties from the physical system of transmission (MacNeilage & Davis, 2000). The differences also suggest that such a program will ultimately be more explanatory than one that assumes that a great deal of phonology is arbitrarily furnished by Universal Grammar. In light of the sign language system, it seems non-explanatory to take for granted that a feature like [coronal] or a constraint like NO CODA is universally generated for all human language. How then to explain a feature like [head] or a constraint like ONE

FINGER GROUP in sign language? This is an allusion to Mandel's (1981) Selected Finger Constraint, which states that only one group of fingers may be selected in a morpheme. The effect of this constraint is also evident within the syllable. Are we endowed with two UG's? This is not likely.

But the similarities between the syllables of signed and spoken languages are significant as well. First, in each modality the syllable organizes lower phonological elements. Second, the syllable is distinguishable from the morpheme and the word, and nonisomorphic with those structures in both modalities. And third, the syllable is in essence a prosodic unit, a unit that is part of the rhythmic system and not part of the lexical system. It is perhaps especially interesting that there is a strong rhythmic effect in sign language in the form of the monosyllable "conspiracy" despite the fact that there is no oscillating mandible to provide a rhythmic frame.

There are many other phonological similarities in the two systems beyond those found in the syllable (Sandler & Lillo-Martin, 2006). For example, both systems have sequential structure (Liddell, 1984; Sandler, 1986), autosegmental structure (Sandler, 1986, 1989), hierarchical organization of phonological features (Corina and Sagey, 1989; Sandler, 1987, 1993a), discrete assimilation rules (Sandler, 1993b), a distinction between lexical (structure-preserving) and postlexical (non-structure-preserving) phonological processes (Padden & Perlmutter, 1987; Sandler, 1993b, 1999a), and a hierarchy of prosodic constituents (Nespor & Sandler, 1999; Wilbur, 1999x). These similarities show that essentially the same cognitive system underlies language in both modalities, and it is necessary for a comprehensive theory of language to account for these similarities as well as differences. These and other similarities also mean that some properties of phonology are **not** directly derivable from the physical system; we must look to higher levels of organization to account for them, as I have argued elsewhere (Sandler, 2006). A theory of the evolution of language must also take this array of discoveries into account.

Bimodal Language and its Origin

The existence of natural language in the manual/visual modality shows that speech does not equal language. Sign language shares key properties with spoken language, including the existence of a phonological level of organization. This phonology is forged from the physical transmission system in tandem with higher level organizing mechanisms. The discovery that humans can "do" language in two different modalities may lead to a variety of conceptions of the nature of the human language capacity. One might assume, for example, that oral and manual language are just different instantiations of the same thing, and that the difference is essentially trivial. However, we have seen that this is not the case. Instead, phonological differences are far-reaching and require theoretical motivation and explanation. An opposing conclusion is that the two modalities are actually so different that they instantiate language in ways that are mutually exclusive. That is, humans have the capacity for two distinct language systems. But this view is also inadequate, as it overlooks two essential properties of these language systems. First, the similarities are also far from trivial. Second, the modalities are not mutually exclusive. Instead, both manual and oral channels are exploited by deaf and hearing alike, in the service of language, an observation that I will expand below. These properties lead to a third theory. The third theory, which can only be painted in broad strokes here, holds that language is essentially bimodal. We evolved to use **both** the hands and the mouth as the vessels of language. Each modality brings a different aspect of a **unified** capacity present in all human linguistic communication (Sandler, 2003).

In his contribution to this volume, Peter MacNeilage proposes that "A theory of the evolution of speech must *begin* with a conception of what it is like now, even though it cannot end there." Extending the notion of "speech" to consider "language transmission" more generally, we see much evidence for bimodalism. First, is the fact that humans are capable of both spoken and sign language. The second piece of evidence comes from manual gestures that universally accompany and supplement speech, the importance of which has

attracted a good deal of attention in recent years (McNeill, 1992, 2002). Hearing children gesture with their hands before they speak. As they begin to acquire spoken words, they first use either a spoken word or a gesture but not together (Goldin-Meadow, 2003). This complementary distribution of speech and gesture in small children suggests that the two perform the same function for them. Only after they begin to develop an explicitly linguistic system does gesture become an auxiliary communicative mode. When it does, this supplementary mode becomes important: hearing people across all cultures augment their spoken language with iconic (and other) manual gestures (McNeill, 1992). A rapidly growing body of research shows that these gestures often add information relevant to the verbal signal, but not present in it. They are part of the message.

Furthermore, bimodalism is bimodal: Deaf people also augment their language with gesture. For them, the primary linguistic signal is made mainly by the hands, and the gestures are made with the mouth. Just as speakers use hand gestures to describe visual properties of referents, signers express visual, tactile, and even auditory impressions with mouth gestures that co-occur with the signed linguistic description (Sandler, 2003). In fact, the mouth is very active during signing, performing a variety of linguistic (non-gestural) functions as well (Boyes-Braem & Sutton-Spence, 2001).

Added to this view of the way things are now is research that hints at an evolutionary precursor to bimodal language, specifically, research on mirror neurons in monkeys (see Fogassi, this volume). Mirror neurons are located in a brain region proposed to be homologous with Broca's area in humans. These neurons discharge when the monkey performs certain actions, and also when the monkey observes the experimenter performing the same action. Rizzolatti, Fogassi, and colleagues (Gallese, Fadiga, Fogassi, & Rizzolatti, 1996) hypothesize that this phenomenon underlies imitation as a learning mechanism used by humans in the translation of perceived phonetic gestures into motor commands in speech. Particularly intriguing in this regard is the discovery that some mirror neurons discharge when either the hand or the mouth moves, but only when the movements have the same goal, i.e., in response to the same

"behavioral meaning" (Gentilucci and Rizzolatti, 1990). Extrapolation tempts the following speculation: there is a neural substrate in primates that links meaning and its physical expression, regardless of whether that expression is oral or manual.

It may well be that the form of the earliest human language was fully bimodal, recruiting both oral and manual expression equally. Oral transmission emerged as primary for hearing individuals at some point in our evolutionary history, but manual expression survives as an option for both the deaf and the hearing, and robust vestiges of a bimodal system remain in both primarily oral and primarily manual modalities, e.g., in the form of co-language gesture.

Conclusion

Comparing the syllable unit of sign language to spoken language reveals differences and similarities in phonological organization across the two modalities. The differences demonstrate clearly that part of phonological organization is linked directly to the physical mode of transmission. This argues for an approach that attempts to derive some of phonology from physical properties of the system and against an approach stipulating a universal pool of formational elements that are generated arbitrarily. The similarities in syllables (and in other aspects of phonology described elsewhere) suggest that some characteristics of phonological organization arise from higher levels of patterning common to both modalities—in the case of syllables, the systemic distinction and interaction between the individual articulatory events involved in producing a meaningful element and its more global prosodic traits. Both kinds of organization result in predictable properties in the phonology of language in each modality. It is an extraordinary fact about humans that we have a natural command of two kinds of phonology, each grounded in a dramatically different physical modality, the oral/aural modality and the manual/visual modality.

In the spirit of the research program to which this volume is dedicated, we must use our conception of the present in order to probe the past. To do this, we focus on four observations that fall out from the discussion presented here.

(1) Phonological similarities between modalities are fully universal across all languages; (2) phonological differences between modalities are fully general across languages in each modality; (3) humans have a natural ability to use both modalities; and (4) language in each physical modality is supplemented by meaningful co-linguistic gesture transmitted in the other. Add to this the intriguing possibility that mirror neurons in other primates that respond equally to hand and mouth actions are precursors to some aspects of language. Taken together, these observations suggest that speech and sign are part of a single bimodal language system well grounded in earlier stages of evolution.

Acknowledgments

I am grateful to Peter MacNeilage for useful discussion and comments on this paper.

References

Bellugi, U. and Fischer, S. (1972). A comparison of signed and spoken language. *Cognition, 1,* 173–200.

Brentari, D., & Poizner, H. (1994). A phonological analysis of a Deaf Parkinsonian signer. *Language and Cognitive Processes, 9(1).* 69–99.

Boyes-Braem, P., and Sutton-Spence, R.l (Eds.). (2001). *The Hands are the Head of the Mouth: The Mouth as Articulator in Sign Languages. International Studies on Sign Language and Communication of the Deaf; 39,* Hamburg: Signum Verlag.

Corina, D. and Sagey, E. (1989). Are phonological hierarchies universal? Evidence from American Sign Language. *Proceedings from Escol, 6,* Columbus, Ohio: Ohio University Press, 73–83.

Coulter, G. (1982). On the nature of ASL as a monosyllabic language. Paper presented at the annual meeting of the Linguistic Society of America. San Diego, CA.

Current Issues in ASL Phonology, Phonetics and Phonology, (2004). Volume 3. 103–129. New York: Academic Press.

Emmorey, K. (2002). Language, *Cognition, and the Brain: Insights from Sign Language Research*. Mahwah, NJ: Lawrence Erlbaum Associates.

Fogassi, Leonardo and Ferrari, Pier Francesco. (2008). Mirror neurons and the evolution of communication and language. In Davis, B. L. & Zajdo, K. (Eds.) *The Syllable in Speech Production*, London: Taylor & Francis.

Gallese V., Fadiga L., Fogassi, L., Rizzolatti G. (1996). Action recognition in the premotor cortex. *Brain, 119 (2):*593–609.

Gentilucci, M. and Rizzolatti G. (1990). Cortical motor control of arm and hand movements. In Goodale, MA (ed.), *Vision and Action: The Control of Grasping.* Norwood NJ: Ablex. 147–62.

Goldin-Meadow, S. (2003). *Hearing Gesture: How our Hands Help us Think.* Cambridge, MA: Belknap Press of Harvard University Press.

Goldsmith, J. (1976). *Autosegmental Phonology.* Doctoral Dissertation. MIT, Cambridge, Mass. [Published 1979. New York: Garland Press]

Halle, M. (1992). Phonological features. *International Encyclopedia of Linguistics, vol 3.* W. Bright (Ed.), Oxford: Oxford University Press. 207–212.

Hockett, C. F. (1960). The origin of speech. *Scientific American*, 203, 88–111.

Klima, E. & Bellugi, U. (1979). *The Signs of Language.* Cambridge, MA: Harvard University Press.

Liddel, S. (1984). THINK and BELIEVE: Sequentiality in American Sign Language. *Language, 60*, 372–392.

Liddell, S. and Johnson, R. (1989). American Sign Language: The phonological base. *Sign Language Studies, 64*, 197–277.

MacNeilage, P. F. & Davis, B. L. (2000). On the origin of internal structure of words. *Science, 288.* 527–531.

MacNeilage, P.F. and Davis, B. L. (2001). Motor mechanisms in speech ontogeny: phylogenetic, neurobiological and linguistic implications. *Current Opinion in Neurobiology, 11,* 696–700.

MacNeilage, P. F. (1998). The frame/content theory of evolution of speech production. *Behavioral and Brain Sciences, 21.* 499–546.

MacNeilage, P. F. (2008). The frame/content theory. In Davis, B. L. & Zajdo, K. (Eds.) *The Syllable in Speech Production*, London: Taylor & Francis.

Mandel, M. (1981). *Phonotactics and Morphophonology in American Sign Language,* Unpublished Doctoral Dissertation, University of California, Berkeley.

McCarthy, J. (1981). A prosodic theory of nonconcatenative morphlogy, *Linguistic Inquiry, 12*, 373–418.

McCarthy, J. and Prince, A. (1986). *Prosodic Morphology.* Unpublished Ms. University of Massachusetts, Amherst, and Rutgers University, New Brunswick, NJ.

McNeill, D. (1992). *Hand and Mind: What Gesture Reveals about Thought.* Chicago: University of Chicago Press.

McNeill, D. (2004). *Language and Gesture.* Cambridge, UK: Cambridge University Press.

Meier, R. (1991). Language acquisition by deaf children. *American Scientist, Vol. 79.* 60–70.

Meier, R. (2002). Why different, why the same? Explaining effects and non-effects of modality upon linguistic structure in sign and speech. In Meier, R. P. Cormier, K., & Quinto-Pozos, D. (Eds.), *Modality and Structure in*

Signed and Spoken Languages, Cambridge: Cambridge University Press 199–223.

Nespor, M. & Sandler, W. 1999. Prosody in Israeli Sign Language. *Language and Speech, 42,* 143–176.

Nettle, D. (1995). Segmental inventory size, word length, and communicative efficiency. *Linguistics, 33.* 359–367.

Newport, E. & Meier, R. (1985). The acquisition of American Sign Language. In D. Slobin (Ed.), *The Cross-Linguistic Study of Language Acquisition, Vol. 1.* 881–938. Hillsdale, NJ: Lawrence Erlbaum Associates.

Padden, C. and Perlmutter, D. (1987). American Sign Language and the architecture of phonological theory. *Natural Langaueg and Linguistic Theory, 5,* 335–375.

Padden, C. (1988). *Interaction of Morphology and Syntax in American Sign Language.* New York: Garland Publishers. [1983: Doctoral dissertation, University of California, San Diego]

Sandler, W. & Lillo-Martin, D. (2002). Natural sign languages. In Aronoff, M. & Rees-Miller, J. (Eds.), *The Handbook of Linguistics.* Oxford: Blackwell. 533–562.

Sandler, W. & Lillo-Martin, D. (2006). *Sign Language and Linguistic Universals,* Cambridge, UK: Cambridge University Press.

Sandler, W. Meir, I., Padden, C., and Aronoff, M. (2005). The emergence of grammar in a new sign language. *Proceedings of the National Academy of Sciences, 102 (7),* 2661–2665.

Sandler, W. (1986). The Spreading Hand Autosegment of American Sign Language, *Sign Language Studies, 50,* 1–28.

Sandler, W. (1987). Assimilation and Feature Hierarchy in ASL, A. Bosch, B. Need, and E. Schiller, (Eds.)., *Chicago Linguistics Society Parasession on Autosegmental Phonology,* 266–278.

Sandler, W. (1989). Phonological Representation of the Sign. Dordrecht: Foris.

Sandler, W. (1993a). Sign language and modularity. *Lingua, 89(4),* 315–351.

Sandler, W. (1993b). Linearization of Phonological Tiers in ASL, in G. Coulter, (Ed.), *Current Issues in ASL Phonology, Phonetics and Phonology Volume 3,* San Diego: Academic Press. 103-129.

Sandler, W. (1999a). Cliticization and Prosodic Words in a Sign Language, in T. Hall and U. Kleinhenz, (Eds.), *Studies on the Phonological Word.* Amsterdam: Benjamins. (*Current Studies in Linguistic Theory*). 223–255.

Sandler, W. (1999b). The Medium and the Message: Prosodic Interpretation of Linguistic Content in Sign Language. *Sign Language and Linguistics, 2(2),* 187–216.

Sandler, W. (2003). On the Complementarity of Signed and Spoken Languages. In Y. Levy and J. Schaeffer (Eds.), *Language Across Populations: Towards a Definition of SLI.* Mahwah, N.J.: Lawrence Erlbaum, 383–409.

Sandler, W. (2004). Phonology, phonetics and the nondominant hand. In L. Goldstein, D. Whalen, and C. Best (es.) *LabPhon 8.* Berlin, New York: De Gruyter.

Senghas, A., Kita, S., & Ozyurek, A. (2004). Children creating core properties of language: Evidence from an emerging sign language in Nicaragua. *Science, Vol 305.* 1779–1782.

Stokoe, W. (1978). [1960]. *Sign Language Structure.* Silver Spring, MD: Linstok Press.

Supalla, T. and Newport, E. (1978) How many seats in a chair? The derivation of nouns and verbs in American Sign Language in P. Siple, (Ed.). 91–133.

Sutton–Spence, R. & Woll, B. (1999). *The linguistics of British Sign Language: An introduction.* Cambridge, England: Cambridge University Press.

van der Kooij, E. (2002). *Phonological Categories in Sign Language of the Netherlands.* Unpublished Doctoral Dissertation. University of Leiden. Utrecht: Holland Institute of Lingustics.

Wilbur, R.B. (1999). Stress in ASL: Empirical evidence and linguistic issues. *Language and Speech, 42,* 229–251.

Woll, B., Sutton-Spence, R., & Elton, F. (2001). Multilingualism: The global approach to sign language. In Lucas, C. (Ed.) *The Sociolinguistics of Sign Languages.* Cambridge, England: Cambridge University Press. 8–32.