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

Wendy Sandler
The University of Haifa

wsandler@research.haifa.ac.il

2008. *Ontogeny and Phylogeny of Syllable Organization, Festschrift in Honor of Peter MacNeilage*, Barbara Davis and Kristine Zajdo (Eds.), New York: Taylor Francis

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

Wendy Sandler
The University of Haifa

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 from it. After a brief introduction to sign language phonology in general (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, that of 'syllable' takes more convincing, which is what Section 4 attempts to do. The sign language syllable is distinguished from other units such as the word and the 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 a 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 such as these underlie phonological differences, providing support for the position of MacNeilage and Davis (2002) 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 does derive from the physical properties of the system and is therefore different in each modality. Instead, we need a theory that explains both commonalities of phonological organization as well as differences in that organization, differences that were chiseled out of the raw material presented by 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 of this kind. In Section 6, 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. Section 7 concludes the chapter.

1. Two kinds of natural language

Sign languages are natural languages that arise spontaneously wherever a group of deaf people has an opportunity to gather and meet regularly (Klima and Bellugi, 1979; Senghas et al 2004; Sandler 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 throughout the world (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, 2005). Certain key areas of the brain are active in the control of spoken and sign language alike (Emmorey, 2002). And sign languages subserve the full range of communicative functions as spoken languages, even including artistic forms such as poetry (Klima and Bellugi, 1979; Sutton-Spence and Woll, 1999; Sandler and Lillo-Martin, 2002, 2005). 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 the other.

2. Sign languages have phonology[1]

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. Despite their iconic and gestural origins, Stokoe showed that there is also a meaningless level of structure in a sign language and, in so doing, he 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., Mandel, 1981; Battison, 1978), and that the form of a word may change in predictable ways in different morphophonological contexts (e.g., Liddell and Johnson, 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 that are 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 phoneme in spoken language. However, he believed that the inter-organization of these elements within a sign is different from that of the spoken word, that the elements occur simultaneously, and not sequentially like the consonants and vowels of a spoken word. Other researchers have found evidence that there is some sequential structure within a sign, although it is more limited than that typically found in spoken words (Liddell, 1984; Sandler, 1989). My own work adopts nonlinear theories of phonology and morphology (e.g.,

[1] While most of the research reported here was conducted on American Sign Language, research on other sign languages such as Israeli Sign Language and Sign Language of the Netherlands has uncovered the same basic phonological structure.

Goldsmith, 1976; McCarthy, 1981) to create a model of sign language structure that reveals both simultaneous and sequential properties (Sandler, 1986; 1989; Sandler and Lillo-Martin, 2005). The analysis of the syllable to be presented in Section 4 assumes this model, which we describe briefly in the following section.

3. Simultaneous and sequential structure in the sign

In a typical sign, the hand, configured 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 (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 (2).

Figure (1) JUST-THEN (ISL)

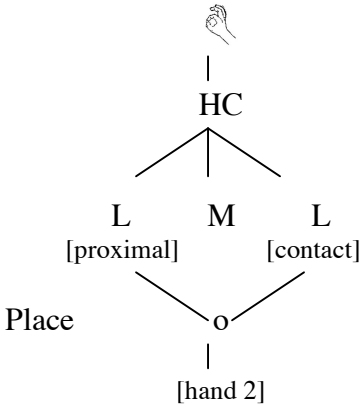
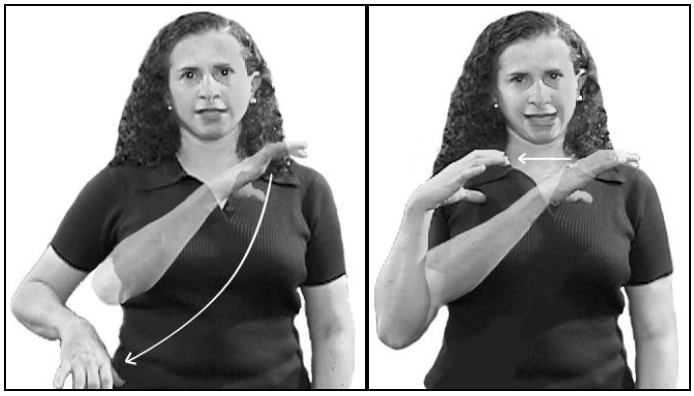


Figure (2) partial representation of JUST-THEN

The model used for the representation is the Hand Tier model (Sandler, 1986, 1989), which is 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 that are distinguished only by one feature in a sequential position within the sign. Like *chap* and *chat* in English, the signs CHIRISTIAN and COMMITTEE are distinguished by the final segment only, as pictured in Figure (3) and illustrated schematically in Figure (4).

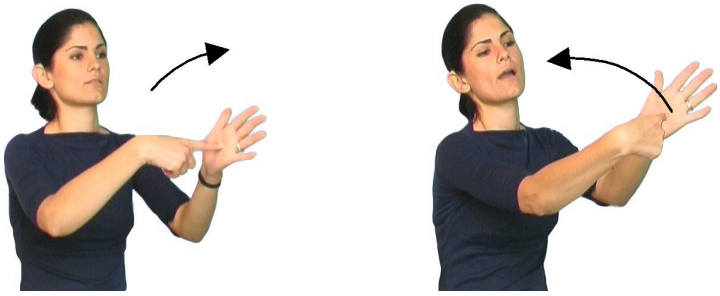


a. CHRISTIAN b. COMMITTEE

Figure (3) minimal pair in ASL distinguished by sequentially occurring features

Figure (4) schematic representation showing features distinguishing CHRISTIAN and COMMITTEE in the final segment of each word

While minimal pairs of this kind, distinguished by sequentially occurring features, are admittedly rare, a good deal 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 location that is 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 (Padden, 1988; Liddell, 1984). Two pictures of the ISL verb SHOW appear in (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 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 (6).



a. I-SHOW-YOU b.YOU-SHOW-ME
Figure (5) verb agreement.

Figure (6) Agreement is marked on the first and last locations

The signs in (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 (1) and (2) above, despite their morphological complexity. All have the canonical structure, LML. As I will explain in the next section, this structure is monosyllabic.

4. The syllable in sign language[2]

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 so labeled bears significant similarity to the syllables of spoken languages. This section will demonstrate that the syllables of sign language are the anchor to which lower meaningless elements are bound, that they are required in order to explain morpho-phonological processes, and that they are prosodic in nature. In all these ways, they are like spoken language syllables. In subsection 4.3, I will describe the differences that also exist between syllables in the two modalities.

We begin with a definition. 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 (1). An example of a sign with handshape change and path movement together is DROP, shown in (9) below.

In order 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 the word. The two sign languages I have studied, ASL and ISL, each have many words that are both monomorphemic and monosyllabic, like JUST-THEN shown in Figure (1). Monomorphemic but disyllabic words like ISL REVENGE, shown in (7) can also be found.

As we have seen in the verb agreement example in Figure (6), a word may consist of several morphemes but still be monosyllabic. And finally, bimorphemic words, like many compounds or

[2] Much of the material in Section 4 summarizes a more detailed treatment in Sandler and Lillo-Martin, 2005.

words with sequential affixes like ISL SEE-SHARP (‘discern by seeing’) shown in Figure (8), may also be disyllabic.

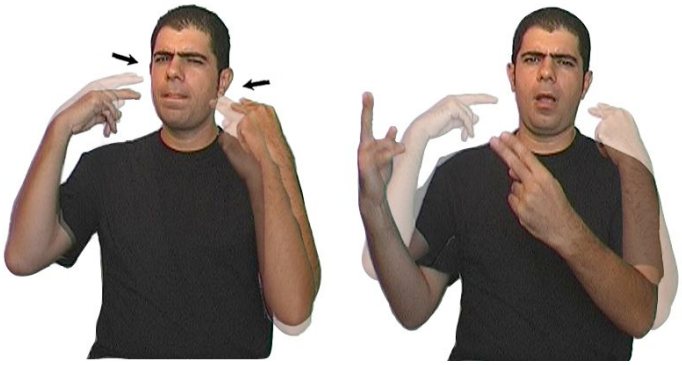


Figure (7). Monomorphemic disyllabic sign: ISL REVENGE

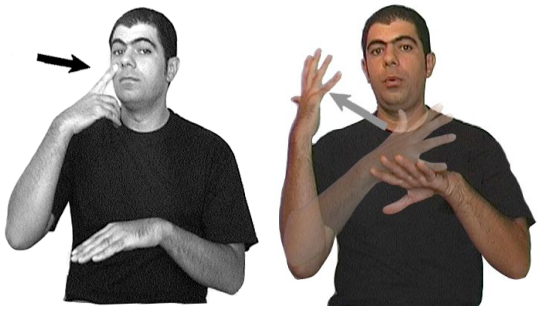


Figure (8). Bimorphemic disyllabic sign. ISL SEE-SHARP (‘to discern by seeing’)

The different relationships between syllables and meaningful units that are found in sign languages are summarized in Table 1. By far the most common kinds of words across sign languages are the first and third shown in bold in Table 1, i.e., words that are monosyllabic regardless of morphological structure.[3]

ω | **μ** | **σ**	ω | μ / \\ σ σ	**ω** / \\ **μ μ** \\ / **σ**	ω / \\ μ μ | | σ σ
monomorphemic monosyllabic words	monomorphemic disyllabic words	**bimorphemic monosyllabic words**	bimorphemic, disyllabic

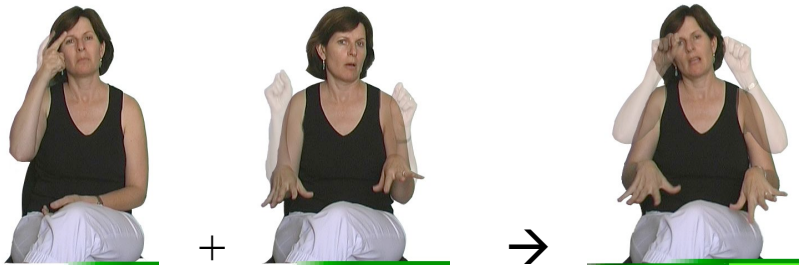
Table 1. The word, the morpheme, and the syllable are distinguished by their cooccurrence patterns. All the possibilities shown are attested, but those in bold are most common.

The relation between the syllable and the word reveals a clear modality effect. In most spoken languages, especially those with morphological complexity, words very often consist of more than one syllable. In sign language, despite the non-isomorphism between the word and the syllable,

[3] Forms that are considered to have more than one syllable for the purposes of Table 1 are only those that have two different syllables; reduplicated forms are not included.

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. An example is one of the ASL compounds for the concept FAINT, formed from MIND+DROP, pictured in Figure (9). MIND and DROP each consist 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 (10).[4]



a. MIND b. DROP c. FAINT

Figure (9). Hand configuration assimilation in an ASL compound.

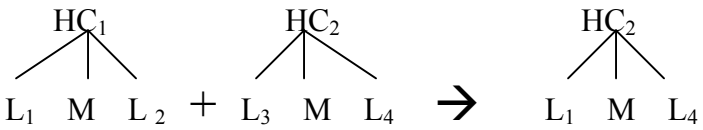


Figure (10). Two syllables reduce to one, producing the canonical, monosyllabic form of a sign (Sandler, 1989, 1999a)

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 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.

4.1. 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 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 have been proposed in research on American Sign Language.

[4] The autosegmental relation between the hand configurations and the locations under compound reduction, shown in Figures (9) and (10) is one of the phenomena that motivated the Hand Tier Model (Sandler, 1987, 1989).

Brentari and Poizner (1994) provide evidence that the syllable is a unit of phonological organization by showing that the 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 (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 that do exist. 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 and Bellugi (1979). ASL nouns that are derived through reduplication have their stress on the first syllable (Supalla and 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 has not reduced to a monosyllable and remains 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, summarized in Table 2, are singling out a prosodic unit, the syllable, and not a morphological or lexical unit. In other words, it is specifically the rhythmic aspect of the syllable unit that is at work in each of these constraints and processes, and rhythmicity is prosodic by definition.

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)

Table 2. Evidence for the syllable in American Sign Language

4.2. 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. So, English clusters that begin with a stop can maximally be followed by one other consonant, which must be a liquid or glide (giving us *proud, plus,* and *puce,* but not **pnack* or **pfack*, for example). 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.[5]

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 and 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.[6]

4.3. 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. The 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.[7] 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, e.g.,  . Second, most of these groups can be configured in one of four different positions. Demonstrated here only with the all-five fingers group, the positions are: open  , closed  , bent  , or curved  . Third, the hand can be positioned in any of several different orientations; two examples are  and  . Finally, the hand can touch or approximate any of a large number of places of articulation on the body.[8] The ASL signs SICK and TOUCH in Figure (11) illustrate just two such places. The Hand Tier model (Sandler, 1989, Sandler and Lillo-Martin, 2005) proposes four major body areas – the head (e.g., Figures (8a),

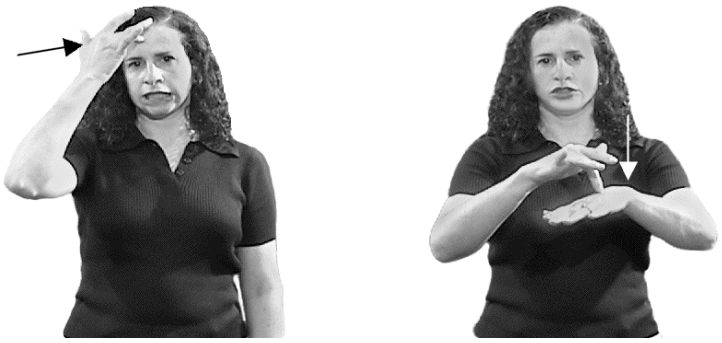
[5] Using a different model of sign phonology from the on assumed here, Brentari (1998) argues further that all phonological elements that are dynamic have the syllable as their domain.

[6] 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 (Nespor and Sandler, 1999), and the intonational phrase in ASL (Wilbur, 1999).

[7] Many signs involve both hands, but I do not deal with this articulatory option here because it does not bear on the present discussion.

[8] For the sake of the discussion, I consider only places of articulation that are in relation to the body, and can therefore be considered system internal, and ignore those places of articulation that are in space. Whether these spatial places are truly linguistic entities is a matter of current controversy (see Sandler and Lillo-Martin, 2005, for discussion).

(10a)), the trunk (e.g., Figure (3)), the nondominant hand (e.g., Figures (1), (10b) and the nondominant arm – and nine more specific 'settings' (such as [hi], [contralateral], etc.) at each of those major areas. Figures (9) above and (11) below illustrate two out of the nine possible different settings on the head, ipsilateral in the sign DROP illustrated in Figure (9), and central in the sign SICK, illustrated in Figure (11).



a. SICK b. TOUCH

Figure 11. Two different places of articulation (ASL)

So even a rough 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 do 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).[9]

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 do have sonority in the form of relative visual salience (e.g., Brentari, 1990, 1998; Perlmutter, 1992; Sandler, 1993),

[9] Such 'tunes' in sign language have been given the label, *superarticulatory arrays* (Sandler, 1999b).

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 and Lillo-Martin 2005 for an explanation). The difficulty in finding a parallel in this regard stems from a fundamental difference in the architecture of the two transmission systems. In spoken language, the source of energy is the lungs, and the relative sonority of the acoustic signal is determined by properties of the filter, the vocal tract. 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.[10]

4.4. Constructing a lexicon: less feature variegation within the sign syllable, but more phonetic features in the system

In Section 3, evidence was presented for sequential structure in the sign. However, the segmental structure of sign language is different from that of 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. I'll illustrate this characteristic with the sign JUST-THEN, pictured in Figure (1) and represented schematically in Figure (2). For clarity, let's start by looking at an SPE-type feature matrix for the English monosyllabic word, *fit*, in Figure (12), and compare it with the feature matrix of ISL JUST-THEN shown in Figure (13).

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.[11] 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 and Fischer, 1972; Emmorey, 2002).

[10] This position, which contrasts in some ways with my own earlier work (Sandler, 1989, 1993), is expanded in Sandler and Lillo-Martin (2005).

[11] 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.

Signs, then, 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 words that are phonologically distinct from one another, and, if that is the case, for 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, we find that 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 (2005) propose that sign languages require 30, a set that is almost twice as large as that of spoken language. Other models of sign language phonology propose even larger numbers of features.[12]

An interpretation of these facts is inspired by work by Nettle (1995), which compared 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.[13]

The explanation is simple, and it 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.

5. The relation between the physical system and phonology

In the previous section, many qualitative and quantitative differences in the nature and organization of the syllable in the two natural language modalities were demonstrated. These 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 in spoken language phonology, a capacity that is counterbalanced by a limited amount of variegation within a syllable.

The very differences between the sign language syllable and the spoken language syllable provide support for MacNeilage and Davis' research program, which seeks to derive phonological properties from the physical system of transmission (MacNeilage and 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 unexplanatory to take it for granted that a feature like [coronal]

[12] 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.

[13] Presumably, all 147 segments in !Xu can be distinguished using Halle's 18 distinctive features.

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?[14] Are we endowed with two UGs? I will argue below that 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 and Lillo-Martin, 2005). For example, both systems have sequential structure (Liddell, 1984; Sandler, 1986), autosegmental structure (Sandler, 1986, 1989), hierarchical organization of phonological features (Sandler, 1987, 1993a; Corina and Sagey, 1989), discrete assimilation rules (Sandler, 1993b), a distinction between lexical (structure-preserving) and postlexical (non-structure-preserving) phonological processes (Padden and Perlmutter, 1987; Sandler, 1993b, 1999a), and a hierarchy of prosodic constituents (Wilbur, 1999; Nespor and Sandler, 1999). 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 the differences. These and other similarities also mean that some properties of phonology are **not** directly derivable from the physical system, that we must look to higher levels of organization to account for them, as I have argued elsewhere (Sandler, in press). A theory of the evolution of language must also take this array of discoveries into account.

6. 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, the phonological differences are far-reaching and require theoretical motivation and explanation. An opposing conclusion that could be drawn 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, as we have seen. Second, the modalities are not, in fact, mutually exclusive. Instead, both manual and oral channels are exploited by all people, 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, and each modality brings to the fore 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, of course, is the fact that humans are capable of both spoken language and sign language. The second piece of evidence comes from manual gestures that

[14] 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.

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, and, as they begin to acquire spoken words, they first use either a spoken word or a gesture but not the two together (Goldin-Meadow, 2003). This complementary distribution of speech and gesture in small children suggests that the two perform the same function for them. It is only after they begin to develop an explicitly linguistic system that gesture becomes an auxiliary communicative mode. When it does, this mode, though supplementary, 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 that is 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 cooccur 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 and 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, and 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.

7. Conclusion

By comparing the syllable unit of sign language to that of spoken language, it has been possible to reveal both 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 which stipulates 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 that are 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 a number of 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) The phonological similarities between modalities are fully universal across all languages; (2) the 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 which 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, one which was fully integrated at an earlier stage of evolution.

References

Bellugi, Ursula and Susan Fischer. 1972. A comparison of signed and spoken language. *Cognition* 1. 173-200.

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

Boyes-Braem, Penny, and Sutton-Spence, Rachel 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, David and Sagey, Elisabeth.1989. Are phonological hierarchies universal? Evidence from American Sign Language. Proceedings from Escol 6. Columbus, Ohio: Ohio University Press. 73-83.

Coulter, Geoffrey. 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, Volume 3. 103-129. Academic Press

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

Fogassi, Leonardo. 2003. Mirror neurons and the evolution of communication and language. **[THIS VOLUME?]**

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, Susan. 2003. *Hearing Gesture: How our Hands Help us Think.* Cambridge, MA: Belknap Press of Harvard University Press.

Goldsmith, John. 1976. *Autosegmental Phonology.* Doctoral Dissertation. MIT, Cxambridge, Mass. [Published 1979. New York: Garland Press]

Halle, Morris. 1992. Phonological features. *International Encyclopedia of Linguistics, vol 3.* W. Bright (ed.), Oxford: Oxford University Press. 207-212.

Klima, Edward & Bellugi, Ursula. 1979. *The Signs of Language.* Cambridge, MA: Harvard University Press.

Liddel, Scott. 1984. THINK and BELIEVE: Sequentiality in American Sign Language. *Language* 60, 372-392.

Liddell, Scott and Robert Johnson. 1989. American Sign Language: The phonological base. *Sign Language Studies* 64, 197-277.

MacNeilage, Peter and Davis, Barbara. 2000. On the origin of internal structure of words. *Science. 288.* 527-531.

MacNeilage, Peter and Davis, Barbara. 2001. Motor mechanisms in speech ontogeny: phylogenetic, neurobiological and linguistic implications. *Current Opinion in Neurobiology. 11.* 696-700.

MacNeilage, Peter. 1998. The frame/content theory of evolution of speech production. *Behavioral and Brain Sciences 21.* 499-546.

MacNeilage, Peter. **[THIS VOLUME?]**

Mandel, Mark. 1981. *Phonotactics and Morphophonology in American Sign Language.* Doctoral Dissertation. University of California, Berkeley.
McCarthy, 1981. A prosodic theory of nonconcatenative morphlogy, *Linguistic Inquiry* 12. 373-418.
McCarthy, John and Prince, Alan. 1986. *Prosodic Morphology.* Ms. University of Massachusetts, Amherst, and Rutgers University, New Brunswick, NJ.
McNeill, David. 1992. *Hand and Mind: What Gesture Reveals about Thought.* Chicago: University of Chicago Press.
McNeill, David. 2004. *Language and Gesture.* Cambridge, UK: Cambridge University Press.
Meier, Richard. 1991. Language acquisition by deaf children. *American Scientist, Vol. 79.* 60-70.
Meier, Richard. 2002. Why different, why the same? Explaining effects and non-effects of modality upon linguistic structure in sign and speech. In Meier, Richard P. Cormier, Kearsy, & Quinto-Pozos, David (Eds.), *Modality and Structure in Signed and Spoken Languages.* Cambridge: Cambridge University Press 199-223
Nespor, Marina and Wendy Sandler. 1999. Prosody in Israeli Sign Language. *Language and Speech* 42: 2&3, 143-176.
Nettle, D. 1995. Segmental inventory size, word length, and communicative efficiency. *Linguistics 33.* 359-367.
Newport, Elissa & Meier, Richard. 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, Carol and Perlmutter, David.1987. American Sign Language and the architecture of phonological theory. *Natural Langaueg and Linguistic Theory 5.i 335-375.*
Padden, Carol. (1988). *Interaction of Morphology and Syntax in American Sign Language.* New York: Garland Publishers. [1983: Doctoral dissertation, University of California, San Diego]
Rizzolatti, Giacomo. 2003. The mirror-neuron system and its relations with language. Paper presented here.
Sandler, Wendy & Lillo-Martin, Diane. 2002. Natural sign languages. In Aronoff, M. & Rees-Miller, J. (eds.), *The Handbook of Linguistics.* Oxford: Blackwell. 533-562.
Sandler, Wendy & Lillo-Martin, Diane. 2005. *Sign Language and Linguistic Universals.* Cambridge, UK: Cambridge University Press.
Sandler, Wendy, Meir, Irit, Padden, Carol, and Aronoff, Mark. 2004 (submitted). The emergence of grammar: Systematic structure in a new language.
Sandler, Wendy. 1986. The Spreading Hand Autosegment of American Sign Language, *Sign Language Studies 50*, 1-28
Sandler, Wendy. 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, Wendy. 1989. *Phonological Representation of the Sign.* Dordrecht: Foris.
Sandler, Wendy. 1993a. Sign language and modularity. *Lingua 89:4* . 315-351.
Sandler, Wendy. 1993b. Linearization of Phonological Tiers in ASL, in G. Coulter, ed.,
Sandler, Wendy. 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, Wendy. 1999b. The Medium and the Message: Prosodic Interpretation of Linguistic Content in Sign Language. *Sign Language and Linguistics* 2:2. 187-216.
Sandler, Wendy. 2003. On the Complementarity of Signed and Spoken Languages. In Y. Levy and J. Schaeffer (eds.), *Language Across Populations: Towards a Definition of SLI.* Lawrence Erlbaum. 383-409.
Sandler, Wendy. In press. Phonology, phonetics and the nondominant hand. In L. Goldstein, D. Whalen, and C. Best (es.) *LabPhon 8.* Berlin, New York: De Gruyter.
Senghas, Anne, Kita, Sotaro, and Ozyurek, Asli. 2004. Children creating core properties of language: Evidence from an emerging sign language in Nicaragua. *Science, Vol 305.* 1779-1782.
Stokoe, William. 1978. [1960]. *Sign Language Structure.* Silver Spring, MD: Linstok Press

Supalla, Ted. and Newport, Elissa. (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, Els. 2002. *Phonological Categories in Sign Language of the Netherlands.* 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 (2&3)*. 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.