21 Sign Language Phonology

DIANE BRENTARI

1 Introduction

In the last edition of this handbook, Goldsmith (1995) wrote in his introduction,

The study of signed languages, such as American Sign Language, promises to have a profound effect on phonological theory, and perhaps ultimately on our understanding of what a human language is. The possibilities that emerge from a linguistic system not constrained by the resources of the vocal tract exploit capacities that had until recently been hidden from linguists' view, and the broadened vista that we have today may in retrospect be as significant for the development of linguistics as was the impact of the Western tradition of the study of non-Indo-European languages.

It is now over 50 years since the first work on sign language phonology appeared (Stokoe 1960). The body of work since 1960 has had three basic aims, which will be referred to by the umbrella terms structure, modality, and iconicity. Under the term structure is included all the work that showed that sign languages were natural languages with demonstrable structure at all levels of the grammar including, of course, phonology. Much progress has been achieved toward the aim of delineating the structures, distribution, and operations in sign language phonology, even though this work is by no means over and debates about the segment, feature hierarchies, contrast, and phonological operations continue. For now, it will suffice to say that it is well-established crosslinguistically that sign languages have structures and hierarchical structures analogous to those of spoken languages. Taken together, the five sign language parameters of handshape, place of articulation (where the sign is made), movement (how the articulators move), orientation (the hands' relation towards the place of articulation), and

© 2011 Blackwell Publishing Ltd. Published 2011 by Blackwell Publishing Ltd.
nonmanual behaviors (what the body and face are doing) function similarly to the cavities, articulators, and features of spoken languages.1 Despite their different content, these groups of features in sign languages are subject to operations that are similar to those of their spoken language counterparts. These broad-based similarities must be seen, however, in the light of important differences due to modality and iconicity effects on the system. Modality addresses the effect of peripheral systems (i.e. visual/gestural vs. auditory/vocal) on the very nature of the phonological system that is generated. Iconicity, which is a specific type of modality effect, refers to the non-arbitrary relationships between form and meaning (Brennan 1990, 2005), particularly visual/spatial iconicity in the case of sign languages.

This chapter will be structured around the three themes of structure, modality, and iconicity because these issues have been studied in sign language phonology (indeed, in sign language linguistics) from the very beginning. Section 2 will outline the phonological structure of signed languages, adding new material to what was covered in Brentari (1995), in the Handbook’s first edition. In Section 3, several typological facts about signed languages regarding word-level phonotactics will be used as an example to illustrate communication modality effects. What is it about the visual nature of signed language that creates this typological niche? This can be answered in terms of signal processing, and experimental evidence from word segmentation supports a modality explanation for the typological facts. Section 4 will focus on iconicity. It will be argued that this concept is not in opposition to arbitrariness, but that iconicity co-exists along with other factors – such as ease of perception and ease of production – that contribute to the inventory of sign language phonological elements. Iconicity is pervasive in the inventory of phonological elements in signed languages, but the distribution of these elements is arbitrary, and it is in the distribution of elements that the true test of phonology lies.

2 Structure

2.1 General Organization of Sublexical Structure

The structure in Figure 21.1 shows the three basic manual parameters – Handshape (HS), Place of Articulation (POA), and Movement (M) – in a hierarchical structure, that of the Prosodic Model (Brentari 1998). This structure presents a fundamental difference between signed and spoken languages. Besides the different featural content, the most striking difference between signed and spoken languages is the hierarchical structure itself – i.e. the root node at the top of the structure is an entire lexeme, a stem, not a consonant- or vowel-like unit.

Both signed and spoken languages have impressive amounts of simultaneous structure, but the representation in Figure 21.1 encodes the fact that a high number of features are specified only once per lexeme in sign languages. This idea will be described in detail below. Since the beginning of the field there has been debate about how much to allow the simultaneous aspects of sub-lexical sign structure to dominate the representation: whether sign languages have the
Figure 21.1 The hierarchical organization of handshape, movement, and place of articulation in a sign, according to the Prosodic Model (Brentari 1998).

The general concept of "root-as-lexeme" present in most phonological models of sign language phonology accurately reflects the fact that sign languages typically specify many features just once per lexeme. These are the inherent features of the Prosodic Model presented in Figure 21.1. Looking at the signs illustrated in Figure 21.2a and their related Place of Articulation and Handshape features in Figure 21.2b, one sees this point clearly. (The major Place of Articulation and Handshape have many feature specifications, all of which remain the same, but for our purposes, it is enough to show them all as a composite symbol in single quotes — e.g., 'L', 'S', etc.). Regarding Place of Articulation, even though it looks like the hand starts and stops in different places in each sign, the major region where the sign is articulated is the same — the torso in WE and SORRY, the horizontal plane in front of the signer in SIT and HAPPEN and the vertical plane in front of the signer in THROW. These are examples of contrastive places of articulation within the system. With regard to Handshape, there is just one Handshape in the first three signs — WE, SORRY, and SIT. The Handshape does not change at all throughout articulation of the sign. In the last sign the two fingers change from closed to open, but the selected fingers used in the handshape do not change. The opening is itself a type of movement, which is described below in more detail. For further elaboration on the separation of selected fingers and aperture, see van der Hulst (1995) and Brentari (1998).

In the space provided here full argumentation for the features and their positions in the structure will not be discussed, but the main point is that relatively more features in sign languages are specified just once per lexeme than is the case with the
Figure 21.2 Illustrations of signs (a) their elements that are specified one time per lexeme (b), their elements that change during the production of the sign (c), and the structure of prosodic features [movement] (d).
features of spoken languages. Tone in tonal languages, and features that harmonize across a lexeme (e.g. vowel features and nasality) behave this way in spoken languages, but fewer features seem to have this type of domain in spoken than in signed languages. And when features do operate this way in spoken languages, it is not universal for all spoken languages. In sign languages a larger number of features operate this way and they do so relatively across sign languages.

The prosodic features in Figure 21.1 are those describing movements within the sign, such as the aperture change just mentioned. These features allow for changes in their values within a single root node (lexeme) while the inherent features do not, and this phonological behavior is the justification for isolating the movement features on a separate autosegmental tier. The trees in Figure 21.2c demonstrate different types of movement features for the signs in Figure 21.2a, and the whole movement feature structure is shown in more detail as the prosodic features (PF) node in Figure 21.2d.

Each specification indicates a different joint (or joints) that can be responsible for a movement – shoulder, elbow, wrist, and hand – progressing top to bottom in the representation from the more proximal joints of the arm to the more distal joints of the hand. In other words, the shoulder articulating the setting movement in WE is located closer to the center of the body than the elbow that articulates a path movement in SORRY and SIT. A sign having an orientation change (e.g. HAPPEN) is articulated by the wrist, a joint that is even further away from the body’s center, and an aperture change (e.g. THROW), is articulated by the hand, furthest away from the center of the body. Notice that it is possible to have two simultaneous types of movement articulated together; the sign THROW has a path movement and an aperture change.

The timing slots projected from the prosodic structure are also shown in Figure 21.2c. The inherent features do not generate timing slots at all, only movement features can do this. When two movement components are articulated simultaneously as in THROW, they align with one another and only two timing slots are projected onto the timing tier. The movement features have been described in detail here because they play an important role in the sign language syllable, discussed in the next section.

2.2 The Syllable

The syllable is as fundamental a unit in signed as it is in spoken languages. One point of nearly complete consensus across models of sign language phonology is that the movements are the nuclei of the syllable. This idea has its origin in the correlation between the function of movements and the function of vowels in spoken languages (Liddell 1984), but this was developed into a theory of syllable structure by Brentari (1990a) and Perlmutter (1992). The arguments for the syllable are based on its importance to the system. They are:

2.2.1 The Babbling Argument Petitto and Marentette (1991) have observed that a sequential dynamic unit formed around a phonological movement appears in
young Deaf children at the same time as hearing children start to produce syllabic babbling. Because the distributional and phonological properties of such units are analogous to the properties usually associated with syllabic babbling, this activity has been referred to as "manual babbling." Like syllabic babbling, manual babbling includes a lot of repetition of the same movement, and also like syllabic babbling, manual babbling makes use of only some of the phonemic units available in a given sign language. The period of manual babbling develops without interruption into the first signs (just as syllabic babbling continues without interruption into the first words in spoken languages). Moreover, manual babbling can be distinguished from excitatory motor hand activity and other communicative gestures by its rhythmic timing, velocity, and spectral frequencies (Petitto 2000).

2.2.2 The Minimal Word Argument This argument is based on the generalization that all well-formed (prosodic) words must contain at least one syllable. In spoken languages a vowel is inserted to insure well-formedness, and in the case of signed languages a movement is inserted for the same reason. Brentari (1990b) observed that ASL signs without a movement in their input, such as the numeral signs 1 to 9 and fingerspelled letters containing no movement in their base form, add an epenthetic movement when used as independent word. Jantunen (2007) observed that the same is true in Finnish Sign Language (FinSL), and Geraci (2009) has observed a similar phenomenon in Italian Sign Language (LIS).

2.2.3 Evidence of a Sonority Hierarchy Many researchers have proposed sonority hierarchies based ‘movement visibility’ (Corina 1990; Perlmutter 1992; Sandler 1993; and Brentari 1993). Such a sonority hierarchy is built into the prosodic features’ structure in Figure 21.2d since movements represented by the more proximal joints higher in the structure are more visible than those articulated by the distal joints represented lower in the structure. For example, movements executed by the wrist are typically larger and more easily seen from further away than those articulated by the hand.

In a study of fingerspelled words used in a series of published ASL lectures on linguistics (Valli and Lucas 1992), Brentari (1994) found that fingerspelled forms containing strings of eight or more handshapes representing the English letters were reduced in a systematic way to forms that contain fewer handshapes. The remaining handshapes are organized around just two movements. This is a type of nativization process by which such forms conform to sign language word-level phonotactics by having no more than two syllables. Crucially, the movements retained were the most visible ones, argued to be the most sonorous ones — e.g. movements made by the wrist were retained while aperture changes produced by the hand were deleted. Figure 21.3 contains an example of this process: the carefully fingerspelled form P-H-O-N-O-L-O-G-Y is reduced to the letters underlined, which are the letters responsible for the two wrist movements.

Some researchers have considered this a manifestation simply of visual loudness' (Crasborn 2001; Sander, and Lillo-Martin 2006). While it is true both in spoken and signed languages that more sonorous elements of the phonology are louder
than less sonorous ones (/a/ is louder than /i/; /l/ is louder than /b/, etc.), the evidence from the nativization of fingerspelled words indicates that sonority has infiltrated the word-level phonotactics of sign languages.

2.2.4 Evidence for Light vs. Heavy Syllables  Further evidence for the syllable comes from a division between those movements that contain just one movement element (features on only one tier of Figure 21.2d are specified), which behave as light syllables (e.g. WE, SORRY, and SIT in Figure 21.2 are light) vs. those that contain more than one simultaneous movement element, which behave as heavy syllables (e.g. THROW in Figure 21.2). It has been observed in both ASL and FinSL that a process of nominalization by movement reduplication can occur only to forms that consist of a light syllable (Brentari 1998; Jantunen 2007). In other words, holding other factors constant, there are signs, such as SIT and THROW, that have two possible forms: a verbal form with the whole sequential movement articulated once, and a nominal form with the whole movement articulated twice (in a restrained manner). The curious fact is that the verb SIT has such a corresponding nominal form (CHAIR), while THROW does not. These facts can be explained by the generalization that the set of forms that allow reduplication have just one simultaneous movement component, and are light syllables, while those that disallow reduplication, such as THROW, have two or more simultaneous movement elements and are therefore heavy.

These sections on word-level phonology and the syllable show clearly that sign languages have all of the elements one might expect to see in a spoken language phonological system, yet their organization and content is somewhat different. What motivates this difference? One might hypothesize that this is in part due to the visual/gestural nature of sign languages, and this topic of modality effects will be taken up in Section 3 of this chapter.

2.3 Prosodic Structure

One area on which sign language phonology has made remarkable progress in the last 15 years has been prosodic structure. Miller (1996), Nespor and Sandler (1999), Sandler (1999), Wilbur (1999), Brentari and Crossley (2002), and Sandler and Lillo-Martin (2006) have worked on various prosodic units, including the
prosodic word (P-word), phonological phrase (P-phrase), and intonational phrase (I-phrase). As in spoken languages, phonological constituents are not isomorphic with their morphological or morpho-syntactic counterparts, but they are related to them in important ways demonstrated for spoken languages in work such as Nespor and Vogel (1986), Selkirk (1984b), and Truckenbrodt (1999). Cues of prosodic structure in signed languages are listed in (1). Prosodic cues are neither all domain cues nor all boundary cues, but domain cues are more prevalent, and more prevalent than domain cues in spoken languages. The P-word cues are all domain cues; in other words, the outputs of Nonmanual Spreading and Handshape Assimilation, discussed below, result in a single value for these properties across the entire P-word. Some of the cues at the level of phrase are domain cues (Non-dominant Handshape Spread) and some are boundary markers (Lengthening). I-Phrase cues are mixed as well; blinks are boundary markers, but many of the nonmanual cues that carry syntactic properties, such as a brow raise over an entire ‘yes/no’ question, are domain cues. Sandler and Lillo-Martin (2006) argued that regardless of the sources of the nonmanual behavior (syntactic, adverbal, etc.) their domains and distribution are prosodic in nature. Nespor and Sandler (1999) have distinguished between the properties of movement, such as those discussed for the P-Phrase, by calling them “rhythm cues” and calling the cues on the face “intonation,” but both rhythm and intonation cues are undoubtedly prosodic. Suffice it to say that many of the important prosodic cues have been noted for several signed languages, but we have yet to understand fully the extent and distribution of all of their uses crosslinguistically.

(1) Prosodic cues in sign languages

a. Prosodic word:
   i. Assimilation of the handshape occurs across a cliticized pronoun (Liddell and Johnson 1989; Sandler 1999).
   ii. Spreading of mouthing (borrowed lip patterns of borrowed spoken words) occurs across two morphological words in a prosodic word (Boyes Braem 1999, 2001; Brentari and Crossley 2002).
   iii. Coalescence of dominant handshape across two morphological words creating one phonological word (Sandler 1999).

b. Phonological phrase:
   i. Spreading of the non-dominant handshape (Nespor and Sandler 1999).

c. Intonational phrase:
   i. Eyeblinks at the right edge of I-Phrases (Wilbur 1994; Nespor and Sandler 1999).
   ii. Changes in leans from left to right in the signing space (Wilbur and Patschke 1998; Boyes Braem 1999).4
   iii. Resetting of all nonmanual behaviors (Nespor and Sandler 1999).

In Figure 21.4 we see examples of markers of each type of constituent. Figure 21.4a, i shows prosodic word markers, using an example from ISL for Handshape...
a. i. P-word: handshape assimilation (ISL)

a. ii. P-word: Non-manual spread (ASL)

b. P-phrase: Non-dominant hand spread (ASL)

c. Blinking at l-Phrase boundaries

Figure 21.4 Prosodic constituent markers in sign languages: (a) Prosodic Word markers in ISL and ASL; (b) Phonological Phrase markers in ASL; (c) Intonational Phrase markers in ASL.
Assimilation – described in (1a, i) of the list above – which appears in forms containing clitics, as well as compounds (Liddell and Johnson 1989; Sandler 1999, Sandler and Lillo-Martin 2006). Notice that the “index finger” handshape, which is used for the first person pronoun when it appears as an independent word, does not appear on the right hand; instead this cliticized form exhibits the same handshape as the following sign READ (the two-finger, “V” handshape); this indicates that the two forms are one P-word. In Figure 21.4a, ii – described in (1a, ii) above – we see an example of Nonmanual Spread (Boyces Braem 1999; Brentari and Crossley 2002). In this case, the mouth posture is the same across the two morphological words COLD and SHOULDER that have become one prosodic word in the expression “cold-shoulder” with the same meaning as the English expression.

Phonological phrase markers are less reliable than P-word or I-Phrase markers, but there are two worth mentioning. In Figure 21.4b is an example of a Handsign Assimilation of the Nondominant hand, described in (1b, i) above. Notice that the left hand remains in the signing space while signing FROM C-O-D-A on the right hand (the ‘c-o-d-a’ is fingerspelled). Notice also that there is a change in the posture of the mouth between the signs FROM and C-O-D-A, so as discussed above, this cannot be judged to be a single prosodic word; if it were, there would be only one mouth posture. This phrase was extracted from the sentence “From a C-O-D-A, [it is a different story].” Brentari and Crossley (2002) found several different uses of H2 spread; it occurs in compounds and idiomatic expressions (such as COLD-SHOULDER, just mentioned) so it was not necessarily the most reliable indicator of a P-phrase. Phrase-final lengthening is slightly more reliable; that is, comparing a phrase-internal and a phrase-final form of the same sign, there will be at least a 1.5 increase in duration in the phrase-final sign. Several different units have been suggested for lengthening. Perlmutter (1992) applied lengthening to the segment, Liddell (1978), Wilbur and Nolen (1986) and Miller (1996) suggested that movement (i.e. syllables) might be the relevant unit. Tang et al. (2010) applied the 1.5 formula to the entire sign.

The three markers of an I-Phrase given in (1c) are all optional; we will discuss only one of these markers – eyblinks. Wilbur (1994) and Nespor and Sandler (1999) have argued that inhibited periodic eyblinks are an I-Phrase boundary marker. These are short blinks that occur either just after or slightly overlapping with the final sign in an I-Phrase. The English translation of the passage in Figure 21.4c is: “[English and ASL are different language]; [WOW! It blows my mind.];” While fairly consistent as an I-Phrase marker across sign languages, there is evidence of crosslinguistic variation in the use of this prosodic cue. Tang et al. (2010) studied blinks in Hong Kong Sign Language (HKSL), Japanese Sign Language (JSL), Swiss German Sign Language (DSGS), and ASL. Across all four signed languages blinks were used to mark the right edges of I-Phrases, just as Wilbur found in ASL, but blinks co-occurred with different cues in JSL than in the other three sign languages. In ASL, DSGS, and HKSL blinks were associated with lengthening, while in JSL, blinks were associated more often with head nods. Even more importantly, it was found that, holding signing rate constant, there
are different blink rates across signed languages, suggesting that even though all four sign languages used blinks to mark I-Phrases, some sign languages may use blinks to mark smaller constituents as well.

Before concluding this section on sign language prosody it is worth mentioning the use of similar prosodic cues in signers and in speakers. There is an undeniable overlap in the content and use of prosodic cues in co-speech gesture and in signed languages, and this topic is receiving attention in the literature. The intentions (precursors to speech acts) of 1-year-old, nonsigning, toddlers at the one-word stage of language acquisition are better understood by caregivers when their gestures and speech cues coincide with one another in timing and content (Balog and Brentari 2008). In terms of timing, there is evidence that pitch accents coincide with manual gestures in adults (Loehr 2004) and that they have an effect on the perceived prominence of accented syllables (Krahmer and Swerts 2007). Furthermore, nonsigning 7.5-month-old infants are more attentive to material in which the prominence expressed visually on the face is in synchrony with the peaks of prosody of the spoken signal; infants at this age are able to segment words from the speech stream better in such a context (Hollich et al. 2005). Finally, several studies have also demonstrated that nonsigners are sensitive to the specific prosodic cues of signed languages. In two studies adult nonsigners were able to perceive the presence of I-Phrase boundaries in a reliable fashion (Fenlon et al. 2007; Brentari et al. 2011); Brentari et al. (2011) also showed a similar result in 9-month old hearing (non-signing) infants.

The use of the prosodic cues mentioned in (1) is clearly arbitrary in sign languages; both distributional evidence and neuro-imaging studies confirm this. Facial expressions that are grammatical are left lateralized in signers but not for non signers (McCullough et al. 2005). Nevertheless, the nature of “gestural competence” in nonsigners can also be informed by the work on sign languages, since there is an overlap in the use of visual/gestural cues as prominence markers and boundary markers in signed and spoken languages.

3 Modality Effects (The Effects of Communication Mode)

The modality effects described here refer to the influence that the phonetics (or communication mode) used in a signed or spoken medium have on the very nature of the phonological system that is generated. How is communication modality expressed in the phonological representation? A few differences between sign and speech that affect the phonetics are: (i) a signed word takes longer to articulate than a spoken one; (ii) an auditory signal is treated differently than a visual signal in processing; (iii) signed words can draw on visual similarities with the entities represented much more easily than spoken words can. I am claiming that strong statistical tendencies in signed and spoken languages (not absolutes) emerge because the communication mode contributes significantly to even the most abstract of phonological structures. Word shape will be used as an example
of how modality effects ultimately become reflected in phonological and morphological representations.

3.1 Word Shape

In this section, the differences in the shape of the canonical word in signed and spoken languages will be described, first in terms of typological characteristics alone, and then in terms of factors due to communication modality. “Canonical word shape” refers to the preferred phonological shape of words in a given language. For an example of such canonical word properties, many languages, including the Bantu language Shona (Myers 1987) and the Austronesian language Yidiny (Dixon 1977), require that all words be composed of binary branching feet. With regard to statistical tendencies at the word level, a preferred canonical word shape is also exhibited by the relationship between the number of syllables and morphemes in a word, and it is here that signed languages differ from spoken languages. Signed words tend to be monosyllabic (Coulter 1982); that is, referring back to the movement structure in Figure 21.2d, the Stokoe et al. (1965) dictionary shows that approximately 83% of the lexical entries are composed of single sequential movements. And, unlike spoken languages, signed languages have a proliferation of monosyllabic, polymorphic words because most affixes in sign languages are feature-sized and are layered simultaneously onto the stem rather than concatenated (see also Aronoff et al. (2005) for a discussion of this point).

This relationship between syllables and morphemes is a hybrid measurement, which is both phonological and morphological in nature, in part due to the shape of stems and in part due to the type of affixal morphology in a given language. A spoken language, such as Hmong, contains words that tend to be monosyllabic and monomorphic with just two syllable positions (CV), but 39 consonants and 13 vowels. The consonant inventory includes as secondary articulations voiced and voiceless nasals, pre- and post-nasalized obstruents, and lateraled obstruents, and the inventory of vowels includes monophthongs and, diphthongs, and seven contrastive tones, including simple and contour tones (Golston and Yang 2001; Andruski and Ratliff 2000). A language, such as West Greenlandic, contains stems of a variety of shapes and a rich system of affixal morphology that lengthens words considerably (Fortescue 1984). In English, stems tend to be polysyllabic, and there is relatively little affixal morphology. In sign languages, words tend to be monosyllabic, even when they are polymorphic. An example of such a form – re-presented from Brentari (1995: 20) – is given in Figure 21.5; this form means “two bent-over upright beings advance-forward carefully side-by-side” and contains at least six morphemes in a single syllable. The agreement forms in Figure 21.7 and all of the classifier constructions in Figure 21.10 (discussed later in Section 4) are monosyllabic as well. There is a large amount of affixal morphology, but most of these affixes are smaller than a segment in size; hence, polysyllabic and monosyllabic words are typically not different in word length. In (2), a chart schematizes the canonical word shape in terms of the number of morphemes and syllables per word.
Figure 21.5  An example of a monosyllabic, polymorphic form in ASL: “two bent-over, upright-beings advance-forward carefully side-by-side.”

(2) Canonical word shape according to the number of syllables and morphemes per word

<table>
<thead>
<tr>
<th>Monosyllabic</th>
<th>Polysyllabic</th>
</tr>
</thead>
<tbody>
<tr>
<td>Monomorphemic</td>
<td>Hmong</td>
</tr>
<tr>
<td>Polymorphemic</td>
<td>sign languages</td>
</tr>
</tbody>
</table>

This typological fact about signed languages has been attributed to communication modality, as a consequence of their visual/gestural nature. This logic predicts that simultaneous systems with a high degree of simultaneous structure both at the level of the stem and in the affixal morphology should be relatively rare in spoken languages and common in sign languages, but notions such as “high degree” and “rare” must be cashed out using available corpora to be viable.

Without a doubt, spoken languages have simultaneous phenomena in phonology and morphophonology such as tone, vowel harmony, nasal harmony, and ablaut marking (e.g., the past preterit in English (sing-pres./sang-preterit, ring-pres./rang-preterit), and even person marking in Hua indicated by the [i]back feature on the vowel (Haiman 1979)). There is also nonconcatenative morphology found in Semitic languages, which is another type of simultaneous phenomenon, where lexical roots and grammatical vocalisms alternate with one another in time. Even collectively, however, this does not approach the degree of simultaneity in signed languages, because, as mentioned in Section 2, many features are specified once per stem to begin with – one handshape, one movement, one place of articulation. Add to that the fact that the morphology is feature-sized and layered onto the same monosyllabic stem, adding additional features but no more linear complexity, and the result is that sign languages have two sources of simultaneity – one phonological and another morphological. I would argue that it is this combination of these two types of simultaneity that causes signed languages to occupy this
typological niche. Many researchers since the 1960s have observed a preference for simultaneity of structure in signed languages, but for this particular typological comparison it was important to have understood the nature of the syllable in signed languages; that is, the syllable is based on the sign’s movement component, as discussed in Section 2.

In the next section experimental evidence is presented showing that the typological fact about signed language word shape is indeed due to the visual/gestural modality, and not accidental or based on language experience.

### 3.2 Simultaneous vs. Sequential Processing

Consider the typological fact just described about canonical word shape from the perspective of the peripheral systems involved and their particular strengths in signal processing. “Simultaneous processing” is a cover term for our ability to process various input types presented roughly at the same time (e.g., pattern recognition, paradigmatic processing in phonological terms); “sequential processing” is our ability to process temporally discrete inputs into temporally discrete events (e.g., ordering and sequencing of objects in time, syntagmatic processing in phonological terms). Despite the fact that many aspects of simultaneous and sequential processing take place in both vision and audition, there are differences in the inherent strengths built into the design of the physiological visual and auditory peripheral systems, as outlined in (3).

| Differences between signal processing in vision and audition (from Brentari 2002) |
|----------------------------------------|------------------|------------------|
| Speed of signal transmission          | Vision           | Audition         |
|                                       | 299, 300 km/sec  | .33 km/sec       |
| Peripheral temporal resolution        | 25–30 ms         | 2 ms             |
| Spatial arrangement information       | Peripheral       | Non-peripheral   |

In general, the advantage in sequential processing goes to audition, while the advantage in simultaneous processing goes to vision. For example, the time required for a subject to detect temporally discrete stimuli is a sequential processing task. The time required for detection in vision vs., audition is different. In vision this period is called the “threshold of flicker fusion” (Chase and Jenner 1993), and in audition the “threshold of temporal resolution” (Kohlrausch, Püschel, and Alpei 1992). Humans can temporally resolve auditory stimuli when they are separated by an interval of only 2 milliseconds (Green 1971; Kohlrausch, Püschel, and Alpei 1992), while the visual system requires at least 20 milliseconds to resolve visual stimuli presented sequentially (Chase and Jenner 1993). Meier (2002) also discusses the ability to judge duration and rate of stimulus presentation; hence the advantage in temporal processing goes to audition.

Regarding simultaneous processing, one effect of the speed of light transmission on the perception of objects is that vision can take advantage of light waves
reflected not only from the target object, but also by secondary reflection from other objects in the environment onto the target object, thereby making use of visual “echo” waves. These secondary reflections are perceived simultaneously with the waves directly reflected from the target object to the retina, enhancing its three-dimensional quality (Bregman 1990). This same echo phenomenon in audition is available to the listener only much later. The result of this effect is that vision allows a more three-dimensional image to be available more quickly due to properties of the signal itself (light vs. sound waves). Moreover, the localization of visual stimuli is registered at the retina and lens, physiologically the most peripheral component of the visual system, while the spatial arrangement of auditory stimuli is resolved at the cortical level and can only be inferred by temporal and intensity differences of the signal between the signals of the two ears (Bregman 1990). Meier (2002) also discusses the transmission property of bandwidth, which is larger in vision, and spatial acuity, which is the ability to accurately pinpoint an object in space (Welch and Warren 1986). All of these factors afford the advantage of spatial resolution to vision.

One might, therefore, expect words in signed and spoken languages to exploit the advantages available to the system.

3.3 Word Segmentation is Grounded in Communication Modality

If the typological difference between words in signed and spoken language described in Section 3.1 is deeply grounded in communication modality it should be evident without language experience. From a psycholinguistic perspective, the phenomenon of word shape can be fruitfully explored using word segmentation tasks, because it can address how language users with different experience handle the same types of items. We discuss two such studies in this section.

The cues that people use to make word segmentation decisions are typically put into conflict with each other in experiments to determine their relative salience to perceivers. Word segmentation judgments in spoken languages are based on (i) the rhythmic properties of metrical feet (syllabic or moraic in nature), (ii) segmental cues, such as the distribution of allophones, and (iii) domain cues, such as the spreading of tone or nasality. Within the word, the first two of these are “linear” or “sequential” in nature, while domain cues are simultaneous in nature and coextensive with the whole word. These cues have been put into conflict in word segmentation experiments, and it has been determined crosslinguistically that rhythmic cues are more salient when put into conflict with domain cues or segmental cues (Vroomen et al. 1998; Jusczyk et al. 1993a, 1999; Houston et al. 2000). Rhythm cues are linear alternations (e.g. ‘chil[dren’, ‘break[ast]), but they unfold more slowly than other linear changes, such as segmental alternations. Rhythm cues also require less specialized knowledge about the grammar; that is, there are only a few prosodic differences that are logically possible if we assume that there is at least one prominent syllable in every word (two-syllable words have three possibilities; three-syllable words have seven possibilities). Segmental
alternations, on the other hand, such as knowing the allophonic form that appears in coda vs. onset position, requires language-particular knowledge at a rather sophisticated level, though infants master it some time between 9 and 12 months of age (Jusczyk et al. 1999).

Word-level phonotactic cues are available for sign languages as well, and these have also been used in word segmentation experiments. This has already been introduced in Section 2, when discussing which features have one value per word (the inherent features) and which ones can change (prosodic features). Some word-level phonotactics are described in (4):

(4) Word-level phonotactics

   a. Handshape: within a word selected finger features do not change; aperture features may change.
   b. Place of articulation: within a word major place of articulation features may not change; setting features (minor place features) within the same major body region may change.
   c. Movement: within a word repetition of movement is possible, or “circle+straight” combinations (“straight-circle”).

The research question is: What properties of sign language play more of a role in sign language word segmentation – the ones that do not change (the simultaneous ones) or the ones that do (i.e. those that are sequential in nature)? These cues were put into conflict with one another in a set of balanced nonsense stimuli that were presented to signers and nonsigners. The use of a sequential cue might be, for example, noticing that the open and closed aperture variants of handshapes are related, and thereby judge a form containing such a change to be one sign. The use of a simultaneous strategy might be, for example, to ignore sequential alternations entirely, and to judge every handshape as a new word. The nonsense forms in Figure 21.6 demonstrate this. If a participant relied on a sequential strategy, Figure 21.6a would be judged as one sign because it has an open and closed variant of the same handshape, and Figure 21.6b would be judged as two signs because it contains two contrastively distinctive handshapes (two different selected finger groups). If, on the other hand, a participant relied on a simultaneous strategy, both of the signs in Figure 21.6 would be judged as two signs.

In these studies, four groups of subjects took part in two experiments. In one study, groups of native users of ASL and English participated (Brentari 2006), and in another study there were groups of native users of ASL, Croatian Sign Language (HZJ), English, and Croatian (Brentari et al. 2011). All were administered the same word segmentation task in which the participants were asked to judge whether controlled strings of nonsense stimuli based on ASL words were one sign or two signs. It was hypothesized that speakers without signing experience might exploit linear cues of the signal since sequential cues, such as strong/weak alternation in foot structure, are so important in spoken languages. It was also hypothesized that that signers would use their language-particular knowledge to judge Figure 21.6a as one sign and 21.6b as two signs. Hypothesis 1 was not confirmed;
there was no significant difference between the signing and non-signing groups’ performance. The strongest result was that both speakers and signers used a domain/simultaneous strategy when segmenting the sign stream into words across all major feature types, but especially in the movement parameter. Hypothesis 2 was confirmed only in the Handshape parameter, and only for signers. Otherwise, both signers and nonsigners used a “1 value = 1 word” strategy overall — a domain strategy — despite the specialized grammatical knowledge for movement and location within words (in the ASL signers’ case).

The conclusion drawn from these word segmentation experiments is that modality plays a powerful role in word segmentation. Domain (simultaneous) cues are stronger in sign languages than sequential cues. Since sequential cues are stronger in spoken languages, it might be reasonable to expect that this might, at least in part, be due to modality as well. This does not mean, however, that a strategy that is dispreferred by the relevant modality is never employed. On the contrary, many spoken languages use domain effects of tone or nasality to signal word boundaries, as has been noted previously, and sequential effects within the word in signed languages have also been noted (see also Aronoff et al. 2005). It does mean that, when faced with a new type of linguistic string, the modality will play a role in segmenting it. Incorporating this factor into the logic of phonological architecture would help explain why certain structures, such as the trochaic foot, may be so powerful a cue to word learning in infants (Jusczyk et al. 1999).

3.4 The Reversal of Segment to Melody

A final modality effect is the organization of melody features to skeletal segments in the hierarchical structure in Figure 21.1. Notice that the content of the
features predicts the skeletal structure in sign languages; as a consequence the features have a higher position and timing slots a lower position in the structure: namely, the reverse of what occurs in spoken languages. Skeletal slots are predictable and hence have a lower position in the structure.9 The reason timing units are higher in the hierarchical structure of spoken languages is because they can be contrastive. Length is not contrastive in any known sign language. But why would this be the case? One reason has already been mentioned in Section 3.2: audition has the advantage over vision in making temporal judgments, so it makes sense that the temporal elements of speech have a powerful and independent role in phonological structure with respect to the melody (i.e. the featural material). One consequence of this might be that the skeletal tier, containing either segments or moras, is more heavily exploited to produce contrast within the system.

In spoken languages, affricates, geminates, long vowels, and diphthongs demonstrate that the number of timing slots must be represented independently from the melody, even if the default case is one timing slot per root node. Examples of the segment-root-feature structure for English are given in (5). A schema for the root-feature-segment structures for both spoken and signed languages is given in (6).10,11

(5) spoken language phonology – root-segment ratios (English)

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>x</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>root</td>
<td>root</td>
<td>root</td>
</tr>
<tr>
<td>d</td>
<td>a</td>
<td>t</td>
</tr>
</tbody>
</table>

(6) Organization of phonological material in signed vs. spoken languages

<table>
<thead>
<tr>
<th>a. Spoken languages</th>
<th>b. Signed languages</th>
</tr>
</thead>
<tbody>
<tr>
<td>root</td>
<td>root</td>
</tr>
<tr>
<td>melody</td>
<td>melody</td>
</tr>
</tbody>
</table>

To conclude this section on modality, researchers working on signed languages confront such issues as the ones described here constantly, since linguistic theories have been developed largely for spoken languages, and as a result we question the influence of the visual modality at every turn. Data from signed languages push the general discussion to an area that is not often considered or possibly simply taken for granted when working on spoken languages alone.
4 Iconicity Effects on Phonological Representation

The topic of iconicity in signed languages is vast, covering all linguistic areas — e.g., pragmatics, lexical organization, phonetics, morphology, the evolution of language — but in this chapter only aspects of iconicity that are specifically relevant to phonological and morphophonemic representation will be discussed in depth. This treatment of iconicity presented here is fresh because research on phonology and research concerning iconicity have been taken up by sub-fields completely independent from one other, one side sometimes even going so far as to deny the importance of the other side. Iconicity has been a serious topic of study in cognitivist, semiotic, and functionist linguistic perspectives, most particularly dealing with productive, metaphorical, and metonymic phenomena (Wilcox 2001; Russo 2005; Taub 2001; Sallandre and Cuxac 2007). In contrast, with some notable exceptions, phonology has been studied within a generative approach, using tools that make as little reference to meaning or iconicity as possible.

"Iconicity" refers to the mapping of a concrete source domain and the linguistic form (Taub 2001); it is one of three Peircean notions of iconicity, indexicality and symbolicity (Peirce 1931[1958]). From the very beginning iconicity has been a major topic of study in sign language research. It is always the "800-lb. gorilla in the room," despite the fact that the phonology can be constructed without it. Stokoe (1960), Battison (1978), Friedman (1976), Klima and Bellugi (1979), Boyes Braem (1981), Sandler (1989), Brentari (1998) and hosts of references cited therein have all established that ASL has a phonological level of representation using exclusively linguistic evidence based on the distribution of forms — examples come from slips of the tongue, minimal pairs, phonological operations, and processes word-formation (see Leuning et al. 2007). In native signers, iconicity has been shown experimentally to play little role in first-language acquisition (Bonvillian 1990) or in language processing in native signers; Poizner, Bellugi, and Tweener (1981) demonstrated that iconicity has no reliable effect on short-term recall of signs; Emmorey et al. (2004) showed specifically that motor-iconicity of signed languages (involving movement) does not alter the neural systems underlying tool and action naming. Thompson, Emmorey, and Golan (2005) have used "tip of the finger" phenomena (i.e. almost — but not quite — being able to recall a sign) to show that the meaning and form of signs are accessed independently, just as they are in spoken languages. Yet iconicity is always there, and every one of these authors mentioned above also acknowledges that iconicity is pervasive.

Iconicity has been dealt with in relative, rather than absolute terms. Frishberg (1975) and Klima and Bellugi (1979) have established that signed languages become "less iconic" over time, but iconicity never reduces to zero and continues to be productive in contemporary signed languages; however, there is no means to quantitatively and absolutely measure just how much iconicity there is in a sign language lexicon. The question, "Iconic to whom, and under what conditions?" is always relevant, so we need to acknowledge that iconicity is generation-specific (signs for TELEPHONE have changed over time, yet both are iconic),
context-specific (the sign for PERIPHERAL is different for a part of the city and for a part of a computer system, yet both are iconic), and language-specific (signs for TREE are different in Danish, Hong Kong, and American Sign Languages, yet all are iconic). Except for a restricted set of cases where entire gestures from the surrounding (hearing) community are incorporated in their entirety into a specific sign language, the iconicity resides in the sub-lexical units, either in classes of features that reside at a class node or in individual features themselves. There may be several layers or types of “resemblance” and not all are appropriately called iconicity (Taub 2001; Russo 2005). The resemblance may be a direct one between the form and the source domain; these are termed “iconic”, or it may be an extension of this primary connection in another domain (“metaphoric”). It may be part of the etymology of a dictionary entry (“frozen”) or it may be latent in the dictionary entry but emerge in particular linguistic contexts (“dynamic”). Finally, iconicity is thought to be one of the factors that makes signed languages look so similar (Guerra 1999; Guerra et al. 2002; Wilcox 2010; Wilbur 2010), and sensitivity to and productive use of dynamic iconicity may be one of the reasons why signers from different language families can communicate with each other so readily after so little time, despite crosslinguistic differences in lexicon and in many instances, also in grammar (Russo 2005). Learning how to use iconicity productively within the grammar is undoubtedly a part of acquiring a sign language.

I will argue that iconicity and phonology are not incompatible. Phonology is both an inventory and the distribution of its elements. Now, after all the work over the last several decades showing indisputably that signed languages have phonology and duality of patterning, one can only conclude it is the distribution (and not the necessarily the inventory) that must be arbitrary and systematic in order for phonology to exist. Iconicity should not be thought of as either a hindrance or as in opposition to a phonological grammar, but rather another mechanism, on a par with ease of production or ease of perception, that contributes to inventories. Saussure was not wrong, but since he based his generalizations on spoken languages, his conclusions are based on tendencies in a communication modality that can only use iconicity on a more limited basis than signed languages.

Iconicity contributes to the phonological shape of forms more in signed than in spoken languages, so much so that we cannot afford to ignore it. I will show that iconicity is a strong initial factor in building signed words, but it is also restricted in outputs, and it can ultimately give rise to arbitrary distribution in the morphology and phonology. Furthermore, I will explain what phonology and phonetics contribute to this interaction, each in its own manner.

4.1 General Effects

It has been shown that the phonological grammar of a sign language can be constructed without the use of iconicity, but what problems can be confronted or insights gained from considering it? It would be odd, even counter-productive, not to use iconicity when it is so readily available. It has been said that signed
languages use iconicity "because they can," since the physical properties of entities, as well as their positions and movements can be quite well represented using a visual-gestural communication modality. For this reason, Mary Brennan (2005) proposed that spoken languages use iconicity in a limited way, not because there is a linguistic restriction against it to which sound symbolism and onomatopoetic forms are the exception, but simply because the sound-speech modality is not suited to it. A lexicon using a sound-based medium simply cannot be constructed based on how entities in the world sound.13

Let us consider the two contexts in which signed languages arise. In most signing communities of the Deaf World, signed languages are passed down from generation to generation not through families, but through communities such as, schools, athletic associations, social clubs, etc. But initially, before there is a community, per se, signs begin to be used through interactions among individuals – either among deaf and hearing individuals ("homesign systems"), or in stable communities in which there is a high incidence deafness. In inventing a homesign system, isolated individuals live within a hearing family or community and devise a method for communicating through gestures that become systematic (Goldin-Meadow 2003). Something similar happens on a larger scale in systems that develop in communities with a high incidence of deafness due to genetic factors, such as what happened on the island of Martha’s Vineyard in the seventeenth century (Groc 1985) and in the case of Al-Sayyid Bedouin Sign Language (ABS; Sandler et al. 2005; Meir et al. 2007; Padden et al. 2010). In both cases, these systems developed at first within a context where being transparent is important in making oneself understood.

Mapping this path from homesign to sign language has become an important research topic since it allows linguists the opportunity to follow the diachronic path of a sign language *al vivo* in a way that is no longer possible for spoken languages. In the case of a pidgin, a group of isolated deaf individuals are brought together to a school for the deaf. Each individual brings to the school a homesign system that, along with other homesign systems, undergoes pidginization and ultimately creolization. This has happened in the development of Nicaraguan Sign Language (NSL; Kegi et al. 1999; Senghas and Coppola 2001). This work to date has largely focused on morphology and syntax, but *when and how does phonology arise in these systems?* Aronoff et al. (2008) have claimed that ABSL, while highly iconic, still has no duality of patterning even though it is about 75 years old. It is well known, however, that in first-language acquisition of spoken languages, infants are statistical learners and phonology is one of the first components to appear (Locke 1995; Aslin et al. 1998; Creel et al. 2004; Jusczyk et al. 1993, 1999).

Confusion between the concepts of ‘transparency’ and ‘iconicity’ often cloud this discussion. Home signer must be transparent to be understood, but signed languages are not transparent; that is, nonsigners cannot guess the meaning of current ASL signs. In one of the first studies of transparency (Klima and Bellugi 1979: 22), 10 hearing participants were shown 90 signs of abstract and concrete nominals, and were reported to be able to make reasonable guesses about meaning for only nine of the 90 signs (10%). Even when the task was multiple-choice,
the participants could not provide correct answers at a level above chance; therefore, we see from Klima and Bellugi’s experiment that ASL is not transparent.

In the next sections we will see examples of iconicity and arbitrariness working in parallel to build words and expressions in signed languages, using the feature classes of movement, orientation, and handshape. The morphophonology of word formation exploits and restricts iconicity at the same time; it is used to build signed words, yet outputs are still very much restricted by the phonological grammar. To make the point that sign languages are much more than their iconic elements, more attention is paid to the small set of forms in sign languages that contradict iconicity (e.g. the sign SLOW signed in its emphatic form surfaces with a faster movement than the base form: Klima and Bellugi 1979), but the more pervasive phenomenon is the use of iconicity and phonology together. The examples below show how iconicity and phonology can be disentangled from one another when both are present.14

4.2 Directional Path Movement and Verb Agreement

One area in sign language grammars where iconicity plays an important role is verb agreement. Salience and stability among arguments may be encoded not only in syntactic terms, but also by visual-spatial means. Moreover, path movements, which are an integral part of these expressions, have a phonological location in the feature tree. Tracing the trajectory from a homesign to a pidgin in a creole system of such forms is spelled out in Senghas (1995) for Nicaraguan Sign Language, and for ABL (Meir et al. 2007; Padden et al. 2010). Here the phonological consequences of this iconic ability will be discussed.

There are three types of verbs attested in signed languages (Padden 1983): those that do not manifest agreement ("plain" verbs), and those that do, which divide further into those known as "spatial" verbs, which take only source-goal agreement, and "agreement" verbs, which take both source-goal agreement, as well as object and potentially subject agreement (Meir 2002; Meir et al. 2007). Meir argues that the main difference between verb agreement in spoken languages and signed languages is that verb agreement in signed languages seems to be thematically, rather than syntactically, determined. Agreement typically involves the representation of phi-features of the NP arguments, and functionally it is a part of the referential system of a language. Typically in spoken languages there is a closer relationship between agreement markers and structural positions in the syntax than between agreement markers and semantic roles; however, sign language verbs can agree not only with themes and agents, but can also agree with their source and goal arguments (Kegl 1985 was the first to note this). The combination of syntactic and semantic motivations for agreement in signed languages was formalized as the "direction of transfer principle" (Brentari 1988), but the analysis of verb agreement as having an iconic source was first proposed in Meir (2002).

Agreement verbs manifest the transfer of entities, either abstract or concrete. The locational-loci of sign language verb agreement are regarded as visual...
manifestations, overt indices of the pronominal elements in question, rather than of grammatical categories such as gender or number (cf. Meir 2002 and references cited therein). Crucially, Meir argues that “DIR,” which is an abstract construct used in a transfer (or directional) verb, is the iconic representation of the semantic notion “path” used in theoretical frameworks, such as Jackendoff (1996: 320); DIR denotes spatial relations. It can appear as an independent verb or as an affix to other verbs. This type of iconicity is rooted in the fact that referents in a signed discourse are tracked both syntactically and visuo-spatially; however, this iconicity is constrained by the phonology. Independently a [direction] feature has been argued for in the phonology, indicating a path moving to or from a particular plane of articulation (Brentari 1988, 1990a, 1998).\(^\text{15}\)

The abstract morpheme DIR and the phonological feature [direction] are distributed arbitrarily both across sign languages (Mathur and Rathmann 2010) and language-internally. In ASL it can surface in the path of the verb or in the orientation; that is, on one or both of these parameters. It is the phonology of the stem that accounts for the distribution of orientation or path as markers in the set of agreement verbs in ASL predicting whether it will surface, and if so, where it will surface. Figure 21.7 provides examples of how the morphological and phonological structures work together, along with diagrams of the [direction] feature involved. In Figure 21.7a we see an example of the agreement verb, APOLOGIZE, that takes neither orientation nor source-goal properties; signs in this set have been argued to have eye gaze agreement with the object (Bahan 1996).\(^\text{16}\) The phonological factor relevant here is that many signs in this set have a distinct place of articulation that is on or near the body. In Figure 21.7b we see an example of an agreement verb that takes only the orientation marker of agreement, SAY-YES; this verb has no path movement in the stem that can be modified in its beginning and ending points (Askins and Perlmutter 1995), but the affixal DIR morpheme is realized on the orientation, palm facing a locus is the object plane.\(^\text{17}\) In Figure 21.7c we have an example of an agreement verb that has a path movement in the stem

![Figure 21.7 Examples of verb agreement in ASL.](http://site.ebrary.com/id/10500910?ppg=729)
4.3 Movement in Event Structure

Wilbur has been involved in research on the relationship between movement and meaning in signed languages since the 1980s (Wilbur et al. 1983, 1999, 2008). Her work has analyzed the prosodic uses of movement for stress, accent, and emphasis as well as the use of movement in aspectual morphology. She recently developed the Event Visibility Hypothesis (Wilbur 2008, 2010), which is a proposal for how the structure of predicates in signed languages adheres to a type of mapping between event structure and phonological form.

In English, event structures are inaccessible via the phonology, although these structures are recoverable through syntactic tests; Wilbur argues that in signed languages, event structure is expressed in the phonology: States (Ss) are [-dynamic] and have no movement; processes (Ps) are [+dynamic] and have a movement; telic events transitions between two non-identical sub-events (P→S), and achievements are transitions between two non-identical sub-events (S→S). Furthermore, processes are homogenous and exhibit no changes other than the passage of time, while telic and inchoative events are heterogeneous. Brentari’s phonological movement inventory (1998) is correlated with the event structure of categories of predicate signs, which are grouped into those that are atelic (states and processes), telic punctual transitions, and telic non-punctual transitions. The semantic, syntactic tests cannot be reiterated in the space allotted here, but evidence is provided in Wilbur (2010) that these distinctions are part of the semantic Aktionsart of the event.

Why is phonology a part of this abstract semantic analysis of event structure? Wilbur argues that by using the features and feature geometry proposed in Brentari (1998) it can be shown that the morphophonology of sign language predicates...
reflects the temporal components of Pustejovsky’s (1995) types of events. These are states, processes, and transitions, which include achievements, and accomplishments). Wilbur further claims that aspectually modified forms, such as resultative, incessant, and continuous, are compositional (each piece contributes meaning), dividing up events into initial, internal, and final temporal sub-periods. Examples include movements for telic and atelic predicates, which exploit the movement features and the associated segmental structure proposed in the Prosodic Model. The features of movement (the prosodic features shown in Figure 21.2d) are used for this analysis. Telic predicates exploit the transitions between the two different specifications for handshape, orientation, setting or [direction] path movements (examples of telic verbs are given in Figure 21.8a–d; this feature was also used in the analysis of verb agreement in the previous section). The feature matrices of the two segments are not identical in this case. Atelic predicates contain a [tracing] or [trilled] movement, which corresponds to phonological shape and manner features in the Prosodic model. Crucially, the segments are identical; there is no change in the feature matrices of the two X slots, only extension in time (examples given in Figure 21.9a–d). Each of the verbs in Figure 21.9 has a [tracing movement], which specifies the shape, and a [trilled] manner feature that indicates that the movement is repeated an uncountable number of times.

This analysis has been able to establish that there is a relationship between event structure and phonology (meaning and form) that is both iconically motivated and phonologically constrained, and Wilbur has argued that these structural components of predicates are one reason why sign languages look so similar to one another.
4.4 Orientation in Handshape in Classifier Constructions is Arbitrarily Distributed

An additional iconic source for a structure that is ultimately distributed arbitrarily involves the orientation of the hand in the handshape of classifier constructions. For our purposes here, classifier constructions can be defined as complex predicates in which movement, handshape, and location are meaningful elements; we focus here on handshape. We will use Engberg-Pederson’s (1993) system, given in (7), which divides the classifier handshapes into four groups. Examples of each are given in Figure 21.10.

(7) Categories of handshape in classifier constructions (Engberg-Pedersen 1993)

a. **Whole entity**: these handshapes refer to whole objects (e.g. “1-HANDSHAPE: person” (Figure 21.10a))

b. **Surface**: these handshapes refer to the physical properties of an object (e.g. “B-B-handshape: flat_surface” (Figure 21.10b))

c. **Limb/body part**: these handshapes refer to the limbs/body parts of an agent (e.g. V-handshape: by_legs (Figure 21.10c))

d. **Handling**: these handshapes refer to how an object is handled or manipulated (e.g. “S-handshape: grasp_gear_shift” (Figure 21.10d))

Benedicto and Brentari (2004) and Brentari (2005) argued that, while all types of classifier constructions use handshape morphologically in a general way, only handshapes in classifier constructions of the handling and limb/body part type
Figure 21.10 Examples of the distribution of phonological and morphological use of orientation in classifier predicates. Whole Entity and Surface/Extension classifier handshapes do not allow morphological use of orientation (top) while Body Part and Handling classifier handshapes do allow morphological use of orientation (bottom).

can use orientation in a morphological way, while whole entity and surface cannot. This is shown in Figure 21.10, which illustrates the variation of the forms using orientation phonologically and morphologically. While the whole entity classifier in Figure 21.10a “person_upside_down” and the surface classifier in Figure 21.10b “flat_surface_upside_down” are not grammatical if the orientation is changed (indicated by an “x” through the ungrammatical forms). The body part classifier in Figure 21.10c “by-legs_be_located_upside_down” and the handling classifier in Figure 21.10d “grasp_gear_shift_from_below” are grammatical when articulated with different orientations.

This analysis requires phonology because the representation of handshape must allow for subclasses of features to function independently with respect to the phonology and morphology of the language according to the type of classifier being used. In all four types of classifiers, part of the orientation specification expresses a relevant handpart’s orientation (palm, fingertips, back of hand, etc.) toward a place of articulation, but only in body part and handling classifiers is it allowed to function morphologically as well as phonologically. It has been shown that these four types of classifiers have different syntactic properties as well (Benedicto and Brentari 2004; Grose et al. 2007).

It would certainly be more iconic to have the orientation expressed uniformly across the different classifier types, but the grammar does not allow this. We therefore have evidence that iconicity is present but constrained in the use of orientation in classifier predicates.
4.5 Conventionalization

A final example of the intertwined nature of iconicity and phonology addresses how a phonological distribution might emerge in sign languages over time (Brentari et al. forthcoming).\textsuperscript{23} Productive handshapes were studied in adult native signers, hearing gesturers (without using their voices), and homesigners in handshapes – particularly, the selected finger features of handshape. Selected fingers indicate which fingers are active in the handshape. The results show that the distribution of selected finger properties is reorganized over time.

We classified handshapes into three levels of selected finger complexity. Low complexity handshapes have the simplest phonological representation (Brentari 1998), are the most frequent handshapes crosslinguistically (Hara 2003; Eccarius and Brentari 2007), and are the earliest handshapes acquired by native signers (Boyes Braem 1981). Medium complexity and High complexity handshapes are defined in structural terms – i.e. the simpler the structure the less complexity it contains. Medium complexity handshapes include one additional elaboration of the representation of a [one]-finger handshape, either by adding a branching structure or an extra association line. High complexity handshapes included all other handshapes. Examples of low and medium complexity handshapes are shown in Figure 21.11.

The selected finger complexity of two types of productive handshapes was analyzed: those representing objects and those representing the handling of objects (corresponding to whole entity and handling handshapes in a sign language). The pattern that appeared in signers and homesigners showed no significant differences: relatively higher finger complexity in object handshapes and lower for handling handshapes (Figure 21.12). The opposite pattern appeared in gesturers, which differed significantly from the other two groups: higher finger complexity in

![Figure 21.11](image)

*Figure 21.11* The three handshapes with low finger complexity and examples of handshapes with medium finger complexity. The parentheses around the B-handshape indicate that it is the default handshape in the system.
handling handshapes and lower in object handshapes. These results indicate that as handshape moves from gesture to homesign and ultimately to a sign language, object handshapes gain finger complexity and handling handshapes lose it relative to their distribution in gesture. In other words, even though all of these handshapes are iconic in all three groups, the features involved in selected fingers are heavily reorganized in sign languages, and the homesigners already display signs of this reorganization.

To summarize this section on iconicity, one can say that each of the elements discussed is iconic and, crucially, also phonological. It has been observed that the co-speech gestures of speakers during narration (see McNeill 2005) contain some of the same surface elements of movement, orientation, and handshape; however, phonology emerges from these properties only when these elements become reorganized (or conventionalized) and assume an arbitrary distribution. Iconicity, like ease of articulation and ease of perception, is a factor that contributes to the phonological inventories of sign languages, and based on the work presented in this section, the distribution of the material is more important for establishing the phonology of signed languages than the source of that material – iconic or otherwise.

5 Conclusion

The more phonologists focus on the physical manifestations of the system – the vocal tract, the hands, the ear, the eyes – signed and spoken language phonology
will look different. The more focus there is on the mind/brain the more sign language and spoken language phonologies will look the same.

This chapter was written in part to answer the following questions: "Why should phonologists, who above all else are fascinated with the way things sound, care about systems without sound?" How does it relate to their interests?" The short answer to those questions is that by looking at the differences and similarities in signed and spoken languages, aspects of work on spoken languages can be seen in a surprising new light, because the range of possibility in expression is considerably broadened, as we see, for example, in the work on visual prosody. Phonologists are in a privileged place to see these differences, because, unlike the case of semantics or syntax, the language medium affects the organization of the phonological system. Using work on signed languages, phonologists can broaden the scope of the discipline to include issues of modality and iconicity, thereby acknowledging that phonology has the potential to exploit a greater range of expressive power than previously thought.

ACKNOWLEDGMENTS

Portions of this chapter have appeared in previous works: Parts of Section 2 and 3.2 are discussed in Brentari (2007). The arguments of Section 3.3 are taken from Benedicto and Brentari (2004) and Brentari (2005). Section 4.3 is drawn from Brentari (2006) and Brentari and Brentari et al. (2011). Portions of Section 4.4 appeared in Brentari (2002).

NOTES

1 The terms Handshape (HS), Place of Articulation (POA), and Movement (M) are used for clarity and ease of exposition here. "Handshape" seems to imply just one hand, but there are also signs that use the arm, or both hands, so "Articulator" is the term sometimes used to cover all of these. "Place of Articulation" is sometimes referred to as "Location" in the sign language literature. And in some models there is no "Movement," but rather "Manner." A structure that includes the non-manual behaviors of the face and body has not been fully worked out, so these are not included in Figure 21.1.

2 I refer the reader to Brentari (1995), Section 1.1, for a history and description of the featural content of the classes of features, known as "parameters" in sign language phonology. Figure 21.1 is a representation from the Prosodic Model (Brentari 1998), which has a particular organization of features and segments, but the point here is only that the root is a lexeme, rather than a C- or V-unit, and this is common to models by Sander (1989) and van der Hulst (1995, 2000), and Channon (2002) as well as the Prosodic Model.

3 See also Jantunen and Takkinen (2010) and Brentari (1995, Section 2.1.2) for more background on the sign language syllable.
4 There is some variation in the constituent claimed to be associated with this cue. It has been documented in Swiss German Sign Language as the P-Phrase (Boyce-Braem 1999), and in ASL as the I-Phrase (Wilbur and Patschke 1998).

5 The dominant hand is the hand used for one-handed signs, and it is the hand on which the more complex handshape appears in two-handed signs. The less complex hand in a two-handed sign is called the nondominant hand (Battison 1978).

6 Not all aspects of sounds are processed cortically; pitch appears to be detected at the level of the brain stem, which is physiologically peripheral with respect to the cortex (Xu et al. 2006).

7 Rhythmic cues are not used at the word level in ASL; they begin to be in evidence at the phrasal level (Miller 1996).

8 These constraints hold for lexemes, but they may be violated in ASL compounds.

9 This point has been addressed similarly in van der Hulst (2000).

10 These are surface representations in English and ASL. In English the /u/ in /dud/ is lengthened before a voiced coda consonant, resulting in an output [dud].

11 [Intensive] forms have a gminated first segment – GOOD vs. GOOD [intensive] ‘very good’, LATE vs. LATE [intensive] ‘very late’, etc. – but no lexical contrast is achieved by segment length.

12 Examples of such gestures that are co-opted “whole-cloth” in ASL are PRAY and SO-SO; in Italian Sign Language (LIS) PERFETTO “perfect,” SOLDI “money,” and COME/PERCHE “why.” No more will be said about this type of grammaticalization. We leave these cases aside; these transparent forms are rare.

13 Iconicity does exist in spoken languages in reduplication (e.g. Haiman 1980) as well as expresives /deophone. See, for example, Bodomo (2006) for a discussion of these in Dagaare, a Gur language of West Africa. See also Okrent (2002), Shintel et al. (2006); Shintel and Nussbaum (2007) for the use of vocal quality, such as length and pitch, in an iconic manner.

14 See also van der Kooij (2002) and Eccarius (2008) for discussions of the interaction between phonology and iconicity.

15 The vertical is one of the three possible planes of articulation in signing space – horizontal, vertical, or mid-sagittal (Brentari 1998).

16 There is debate about exactly what role eye gaze plays in the agreement system, but that it plays a role is not controversial (Neidle et al. 2000; Thompson and Emmorey 2006).

17 There is an emphatic form that has a path movement, added but this is not the form typically used.

18 In many other sign languages, the verb HELP is articulated differently, and in Italian Sign Language, German Sign Language, and Israeli Sign Language HELP has both orientation and path markers for agreement.

19 In ASL we have found that the V-handshape “by-legs” can function as either a body or whole entity classifier.

20 Orientation differences in whole entity classifiers are shown by signing the basic form, and then sequentially adding a movement to that form to indicate a change in orientation.

21 This work is being carried out thanks to NSF grants BCS 0112391 and BCS 0547554 to Brentari, and is being carried out collaboratively with Susan Goldin-Meadow, Marie Coppola, Laura Mazzoni, along with consultants Roberto Ajello and Virginia Volterra.