


The Importance of Autosegmental Representations for Sign Language Phonology

Diane Brentari

6.1. Introduction

In the ideal scenario, evidence from data motivates theory, and, subsequently, the theory makes testable predictions, leading to evidence that supports or challenges the original theory. There is no doubt, however, that our perception of new data is colored by the theoretical lens through which we view it. Two central aspects of Goldsmith’s work have fundamentally changed the lens through which we see phonological data, especially data from sign languages, which is the focus here. This chapter addresses two arenas where sign language phonology was inspired by topics that Goldsmith put on the map, either single-handedly or with others. First, the autosegmental tier allowed the field to expand phonological representations in ways that were more simultaneous and multidimensional in nature than ever before. This leap expanded the general conception of phonology to better accommodate sign language phonology, allowing signed and spoken language phonology to find some common ground within the field. Second, because the autosegmental tier was autonomous from segments...
and intimately linked to the field of prosodic analysis starting with the 
syllable, it paved the way for advances in the formalisms that express 
prosodic generalizations. The syllable was largely ignored in Chomsky 
and Halle’s (1968) *The Sound Patterns of English* (SPE) phonology (but see 
Hooper 1972), and autosegmental phonology was largely responsible for 
resuscitating it.

Since phonological theory should aim to describe the abstract organ‐
ization of the meaningless units of language, it is implicit that it should 
adequately describe the facts and make predictions about the organ‐
ization of both signed and spoken phonological systems. Discussions 
between phonologists working on both types of languages allow for the 
possibility to consider aspects of phonology that might be universal, 
rather than dependent on the modality of transmission. Autosegmental 
phonology created two paths for such interchanges with the construct 
of autosegmental tier and subsequent work on the syllable. The auto‐
segmental tier provided a means to capture, for the first time, the sign 
language “parameters” of handshape, movement, location, and orienta‐
tion, rendering them more comprehensible to phonologists working on 
spoken languages. Since the autosegmental tier is an explanatory tool for 
both sign and spoken languages, it adds to its permanence within the 
field. This is described in section 6.2. In addition, developments regarding 
the syllable that grew out of autosegmental phonology allowed for syl‐
lable weight and sonority to be assessed simultaneously or sequentially 
within the syllable, which are essential when analyzing sign language 
data. The discussion of these issues in section 6.3 suggests that the so‐
nority sequencing principle (SSP) is modality specific (i.e., relevant only 
to spoken languages), while the general concept of sonority applies to 
both signed and spoken languages in assigning and determining syllable 
weight.

6.2. Multidimensionality in Representation

Often, when Goldsmith talked about his own conceptualization of 
autosegmental phonology he produced a co-speech gesture showing a 
three-dimensional dynamic space rotating on an axis similar to one of 
Alexander Calder’s mobiles. This gesture communicates in an optimal 
way what Goldsmith was seeking to reveal about phonological represen‐
tation in a multidimensional space. ¹ Multidimensional images of phono‐
logical features, such as those in figure 6.1 from Clements (1985), clearly 
made visible that the notion of phonemes as beads on a string was in‐
6.1 Multidimensional representations of features. (From Clements 1985; Clements and Hume 1995.)
sufficient to capture a wide range phonological phenomena (Goldsmith 1976, 1985, 1990, 1992; Clements and Goldsmith 1984). Such representations were an essential step in allowing sign language phonologists to conceptualize the sublexical elements of those languages, as will be described in section 6.2.1.

Without explicitly saying so, autosegmental phonology and feature geometry (Clements, 1985; Sagey, 1986; Clements and Hume, 1995) are sometimes invoked as if they were a single theory. Part of this has to do with the dual nature of association lines. Association lines serve two purposes: first, to encode patterns of temporal alignment and coordination among elements in phonological representations; and second, to group elements into constituents that function as single units in phonological rules. Because they are used in both feature trees and autosegmental phonological rules, their distinct theoretical roles in these two theories can be confused. Features produced by a single articulator—that is, the glottis—can be grouped together using an articulatory-based account of feature geometry, but they can have different autosegmental roles. For instance, the features of both tone and voicing are produced by the vocal folds, so features of tone ([high], [low]) and voicing ([spread glottis], [constricted glottis]) could be placed together physiologically; however, since tone is autosegmental in some languages, it will be assigned to an autosegmental tier, but [spread glottis] and [constricted glottis] will not. As we will see handshape and place of articulation (POA) behave in a similar fashion in sign languages.

In Goldsmith (1976), the criteria for establishing autosegmental tiers are phonological and not anatomical or physiological; they include stability, many-to-one association, morphological identity, and long-distance effects. One might think that autosegments are simply a way of representing nonlocal dependencies of all kinds—those that are based on physiology and those that are not, but autosegmental tiers have more flexibility and perform more organizational work in a phonological representation than features do. Goldsmith (1976: 13; cf. Ladd, chapter 5 of this volume) argued for a distinction “between a rule of feature-assimilation . . . and true spreading of a suprasegmental [autosegmental] element,” and that it is “a feature’s behavior with respect to phonological rules that gives it away as suprasegmental, not any phonetic facts.” While analyses of the mechanics of articulation and phonetic spreading have a place in phonological theory, the goal should be to arrive at the underlying principles that govern these sensorimotor activities.
6.2.1. Multidimensionality in Spoken Languages: A Brief Sketch

The prototypical cases of autosegmental spreading are not simply cases of assimilation due to the mechanics of the articulators, but rather are factors about their organization according to specific phonological and morphological domains. Tone is the most obvious case. As Goldsmith and Mpiranya (2011) stated, “one particular characteristic of tone that keeps it distinct from other aspects of phonology is tone’s tendency to shift its point of realization among a word’s syllables based on a global metrical structure which is erected on the entire word.” In this volume, Hyman (chapter 3) and Odden (chapter 4) provide excellent examples of this. Nasal harmony (Piggott 1989), vowel harmony (van der Hulst and van de Weijer 1995; Henderson 1985), and more recently clicks in the Khoisan languages (Bradfield 2014) have also shown characteristics of autosegmental spreading rather than generic feature spreading. The articulatory mechanics of feature spreading can be used to describe how the articulatory gestures are coordinated as a gestural score in spoken and sign languages (Browman and Goldstein 1989; Tyrone et al. 2010), but they do not account for other important aspects of these operations; for example, why feature spreading stops or is blocked when there is no opposing gesture present. For example, the tongue tip can move quite quickly and quite easily, and the velum moves more slowly, and these facts affect assimilation phenomena, to be sure, but to explain other facts of assimilation, such as those of the blocking of voice assimilation in Ren­ daku phenomena in Japanese via Lyman’s law (Itô and Mester 1995), the nature of feature specifications (underlying versus redundant) as well as phonological and prosodic constituents are needed, including the autosegmental tier and the phonological word, among others (see also Piggott 2000; Goldsmith and Mpiranya 2011).

In addition, as Ladd (2014; see chapter 5 of this volume) pointed out, an autosegmental tier combines characteristics of both segments and features. The autosegment was intended to be an entity whose manifestation in the phonetics favor treating it as a feature—an attribute of some segment—but whose function in the phonology suggests that it is a segment in its own right. Treating autosegmental assimilation and feature spreading as the same phenomenon removes any motivation for positing autosegments as an innovative phonological unit with its own criteria. Ladd pointed out that the utility (the genius) of autosegments is that they have this dual character as unordered units with respect to the segment (Merkmale) and as properties (Eigenschaften) that are ordered
at some level of the representation. These are characteristics of features that can be traced to Trubetzkoy (1939).

In sections 6.2.2 and 6.3.2 I will outline ways that sign language data have validated these two roles of features as unordered (Merkmale) and ordered (Eigenschaften) units within the word. I claim that sign languages have a greater capacity than spoken languages do for having independent articulators active at one time. As such, in sign languages there are more possibilities for unordered phonological units, and it may be the case that they depend on autosegmental tiers even more heavily than spoken languages do, validating the constructs of autosegmental phonology as Goldsmith intended.

6.2.2. Multidimensionality in Sign Languages

Sign language phonology has developed very rapidly since its inception in 1960, but in general, theoretical developments proposed for spoken languages do not get tested on sign language data immediately. I will illustrate the development of the field using the American Sign Language (ASL) sign INFORM (figure 6.2) as an example and will consider three of the five parameters of a sign—handshape, POA (location), and movement. The remaining two parameters of orientation and nonmanual behaviors will not be discussed. Handshape concerns the particular fingers that are used and the way the joints of the wrist and hand are configured. In the sign INFORM, all of the fingers are bent at the knuckle joint and the fingertips make contact with the forehead. Place of articulation is where the sign is articulated on the body or in front of the signer; INFORM starts at the forehead and ends in the neutral space in front of the signer. Movement is the dynamic part of the sign characterized by a particular direction or shape of movement as the hands from one place to another. INFORM is articulated with a direction of movement away from the signer’s forehead and an opening gesture of the hand.

The earliest model of sign language phonology proposed by Stokoe (1960; sometimes called the cheremic model) emphasized the unordered, simultaneous nature of the three parameters. He treated handshape, movement, and POA as phonemes in the structuralist sense of the term (Bloomfield 1933). Like spoken language models in the 1950s, Stokoe (1960) focused on providing evidence for features using phonemic contrast, as was practiced in a phonemic account (Bloomfield 1933), but Stokoe was unsure about how to label the bundles of features that made up the three sign language parameters, because phonemes in spoken languages occur one after the other, while sign language par-
Parameters are articulated at the same time. He initially suggested a completely new vocabulary for sign language units (i.e., cheremes), but this was ultimately considered to be counterproductive to the enterprise of promoting dialogue between phonologists working on signed and spoken languages.

Stokoe obtained his PhD in 1941, and during the 1940s and 1950s, the interest in sign languages came primarily from the discipline of anthropology, since Plains Indian Sign Language was of interest to several students of Franz Boas (see Kroeber 1958; Voegelin 1958; West 1960; a review of this work in the field of anthropology appears in Davis 2010). A point of contact between the formal advances made in spoken language phonology using a pre-generative framework (Jakobson, Fant, and Halle, 1952; Jakobson and Halle, 1956) and those in sign language phonology did not occur at this crucial juncture. Moreover, phonologists responsible for pointing out the weaknesses in SPE for a wide range of phenomena—including tone, metrical structure, and prosodic structure—were not aware that SPE also did a poor job of accounting for sign language facts. By the time Stokoe’s work became widely known at the beginning of the 1970s, spoken language phonology was deeply immersed in SPE phonology, and one of the most common questions in the early years of working with Stokoe’s model was whether the sign language parameters were features or segments; in other words, How did they fit into a SPE framework?

6.2 The ASL sign INFORM with its phonological description for handshape, place of articulation, and movement. (From Brentari 2016.)
In 20-20 hindsight we can see that the parameters of handshape, movement, and POA are autosegmental tiers, but the evidence for them was not proposed until almost sixteen years after Stokoe’s initial 1960 publication (Sandler 1986). Additional work in the Stokoe tradition did a great deal to elaborate on the set of features responsible for contrast in ASL, but not on their hierarchical organization (Frishberg 1975; Klima and Bellugi 1979; Lane, Boyes Braem, and Bellugi 1976; Friedman 1977; Battison 1978; Mandel 1981; Poizner 1983).

Four subsequent models of phonological representation will be discussed in more detail (figure 6.3): the hold-movement model, the hand tier model, the dependency model, and the prosodic model. The hold-movement model (figure 6.3a) was largely inspired by SPE (Liddell 1984; Liddell and Johnson 1989). This work advanced the field of sign language phonology, and these authors proposed for the first time an organization of the phonological features of a sign. It was based on and biased toward a sequential representation of ordered segments. The hold-movement model answered the question “how did they fit into a SPE framework?” by proposing temporally ordered segments, within which in all features in each segment were bundled together. Holds (i.e., static elements) were compared to consonants, and movements (i.e., dynamic elements) were compared to vowels.

Within each segment, the features were implicitly organized by parameter, similar to the way the place, manner, and voicing features were implicitly placed together in an SPE feature matrix for a speech segment. Regarding the sign INFORM in figure 6.2 this meant that there was a shift from the simultaneous, unordered treatment of units in the cheremic model to sequential treatment of units in the hold-movement model: the temporal change as the hand moves from the hold at the forehead to hold in the space in front of the signer, and the movement in between (H-M-H)—three segments in total. By the time the hold-movement model became widely accepted within the sign language community in the 1980s, however, spoken language phonologists had moved on and were working out the details of autosegmental phonology and feature geometry in the ways described in section 6.1. There was still a substantial difference between how signed and spoken language phonology was represented.

The hold-movement model captured some important facts about assimilation and moved the field forward; however, the features associated within each segment in the hold-movement model contained a substantial amount of redundant information when compared to their spoken language counterparts—consonants (Cs) and vowels (Vs).
6.3 Four sign language models of phonological representation: (a) the hold-movement model, (b) the hand tier model, (c) the dependency model, and (d) the prosodic model.
hand tier model (Sandler 1986, 1989; Sandler and Lillo-Martín 2006) (figure 6.3b) would be the first to address some of the shortcomings of the hold-movement model by representing handshape as an autosegmental tier, thus partially reanimating Stokoe’s original insights about the simultaneous nature of sign structure. Evidence for handshape as an autosegment is provided by the handshape sequence constraint (Sandler 1986), which formalized the generalization, first observed by Mandel (1981), that a sign only has only one set of selected fingers. Selected fingers are the features that typically denote which fingers can have contact with the body in a sign (note that all of the fingers contact the forehead at the start of the sign). In a sign, such as (figure 6.2), there is a change in the handshape from closed \( \overline{\text{a}} \) to open \( \overline{\text{b}} \); however, all four of the fingers plus the thumb are selected throughout the sign. The [closed] to [open] movement of the fingers is expressed by aperture features. A change in aperture such as this \( \overline{\text{a}} > \overline{\text{b}} \) is permitted in a monomorphemic sign, while a change in selected fingers such as this \( * \overline{\text{a}} > \overline{\text{b}} \) or \( * \overline{\text{b}} > \overline{\text{a}} \) is not permitted in a monomorphemic sign. The relevant domain for the handshape sequencing constraint has been proposed to be the morpheme (Sandler 1986) or the phonological word (Brentari 1998), since historical change progresses in this direction. As compounds containing stems with two different sets of selected fingers become more integrated into the phonology as single words, one of the handshapes assimilates to the other—for example, \( \overline{\text{red}} ^ {\text{slice}} \) has become \( \overline{\text{red}} ^ {\text{slice}} \) (tomato). Crucially, the domain of selected finger spreading is confined to the word; the phonetic facts have nothing to do with this.

The handshape parameter thus has a dual role with regard to selected fingers features, which are autosegmental and unordered within the word, and aperture features, which are segmental and ordered, just as the laryngeal features have a dual role in tone languages: tone (autosegmental, unordered) and voice (segmental, ordered). The sign language parameter of location (POA) also has this dual role in phonology. POA consists of major POA—four areas of the body where a sign is produced: the torso, head, arm, or nondominant hand. These features are autosegmental and unordered (like selected fingers) and appear only once per monomorphemic sign. Within a monomorphemic sign, however, setting features such as [top], [bottom], [ipsilateral], or [contralateral] are allowed to change in an ordered fashion, and are thus segmental and ordered (like aperture).

The dependency model (van der Hulst 1993; van der Kooij 2002; van der Kooij and van der Hulst 2005) (figure 6.3c) and the prosodic model (Brentari 1998) (figure 6.3d) unite both the sequential (ordered) and simultaneous (unordered) nature of signs in their respective representa-
tions even further than the hand tier model does. In both the dependency and prosodic models, the X-slots (timing slots) are demoted from the “root node” status they have in spoken languages; they are derived from features in the underlying form that generates them (van der Hulst 2000; Brentari 2002). Unlike segmental structure in spoken languages, which can be contrastive—in singleton versus geminate consonants or long versus short vowels—length is not contrastive in the sign languages that have been documented thus far. Although these two models differ in important details, both ascribe autosegmental status to features of selected fingers and to major POA because they meet the phonological criteria established by Goldsmith (1976).

One important difference between the dependency and prosodic models is the principle guiding the organization of features. For inform, the two respective models’ representations would look like those in figures 6.3c and 6.3d. In the dependency model, physiology is the major guide in organizing features; both selected fingers (autosegmental, unordered) and aperture (segmental, ordered) are dominated by the same class node in the tree, despite their different phonological roles, because they are handshape features. Major POA and setting features are united for this reason as well. In contrast, in the prosodic model, the phonological behavior is the primary guide in organizing features. All unordered, autosegmental features are located together in the tree in the inherent features branch (i.e., selected fingers, major POA), and inherent features do not generate segments. Ordered features that generate, such as aperture and setting features are found in a different place in the structure, called the prosodic features because ultimately these features also generate syllable nuclei.

Presenting more specific arguments for one or the other of these models is beyond the scope of this chapter, but there are three points to take away from the discussion thus far. First, autosegmental phonology is absolutely essential to signed and spoken language phonology. The domain of selected fingers and major POA are determined by phonological constituency, and not by phonetic facts. The impact of the notion of the autosegmental tier and its properties helped to initiate a great deal of work demonstrating that sign language phonology did not require theoretical tools that were different from those of spoken languages. Second, as Ladd (chapter 5 of this volume) points out, whether a property is ordered or unordered with respect to segments is a crucial distinction for phonology regardless of whether it concerns a spoken or signed language. Finally, since middle of the 1990s it has no longer been in doubt (at least to phonologists) that sign languages have important data to
contribute to the discussion of the abstract categories that constitute phonology. In order to describe the full range of extant phonological systems, theories must strive to handle both types of data. Work by Hale and Reiss (2008) has gone so far as to propose “substance-free” phonology. Although we are still a very long way from achieving that goal, the organization of sign language phonology confirms and validates the need for autosegmental tiers.

6.3. Sonority and Syllable Weight

The study of the syllable was a natural development of autosegmental phonology. The connections between prosodies and autosegmental tiers have been well documented (Firth 1948; see also Coleman, chapter 1 of this volume; Leben, chapter 2 of this volume). Three of the primary motivators of this development were that (i) tone-bearing units were syllabic rather than segmental units (Goldsmith 1976), (ii) the syllable more elegantly accounted for a wide variety of phenomena than the segment in SPE (Hooper 1972), and (iii) segmentation itself was often determined by the role the feature bundle played in the syllable (e.g., vowel-glide alternations; Dell and Elmedlaoui 1985; Levin 1985; Itô 1986). Goldsmith and Larson (1990) also contributed the insight that there are two different levels of sonority—inherent sonority based on a segment’s underlying form and derived sonority (see also Wiltshire, chapter 8 of this volume), which also considers the segment’s context. Spoken languages employ a sonority hierarchy, such as the one in (1), and we will see that a sonority hierarchy is employed in sign languages as well.

(1) Sonority hierarchy in spoken languages (Dell and Elmedlaoui 1985)
low V > high V > glide > liquid > nasal > voiced fricative > voiceless fricative > voice stop > voiceless stop

In addition to assigning syllable peaks, sonority can also important in both signed and spoken languages for assigning syllable weight (Gordon 2006, Gordon et al., 2008). Moraic (or weight) units are employed to assign syllable weight—one mora results a light syllable, while two or more morae results in a heavy syllable—and heavy syllables are more likely to be stressed in many languages (Trubetzkoy 1939; Hayes 1989). In the following sections, two lines of research about the syllable that have been pursued for spoken languages will be discussed as they
pertain to sign languages—(i) the correlation of energy with sonority and (ii) the contribution of simultaneous elements in the syllable to weight—both of which Goldsmith’s early work on the syllable inspired (Goldsmith 1992).

6.3.1. Sonority and Syllable Weight in Spoken Languages: A Brief Sketch

Sonority and syllable weight are related in important ways, which led Gordon et al. (2008) to propose that, particularly in languages that treat CVC syllables as light, both sonority and syllable weight are determined by the amount of energy used to produce the form. In Gordon (2008), for example, the perceptual energy in rimes calculated using Cricket, a custom-made software designed to map an acoustic signal to one that more accurately reflects the perception of that signal by the human auditory system, showed a number of important results. In particular, a rime that ends in a sonorant or voiced coda has greater energy than one closed by an obstruent or a voiceless coda, respectively, particularly in languages that treat CVC as light. The correlation of energy expenditure is important for associating high energy generally with sonority peaks, and with higher syllable weight as well, in both signed and spoken languages.

6.3.1.1. Energy and Sonority in Spoken Languages: The SSP

Evidence for the importance of sonority in phonology also comes from the SSP, which captures the fact that syllable onsets typically increase in sonority from one segment to the next and codas typically decrease in sonority from one segment to the next. The SSP has been shown to be important for determining syllable well-formedness both theoretically (Clements 1990) and experimentally (Berent 2013). Unless there is an L-specific pattern allowing violations—for example, Polish (Rubach and Booij 1990)—even languages that do not allow consonant clusters, such as Korean, show evidence of the SSP. Berent (2013) argued that the SSP is a core property of phonology because Korean speakers are more likely to perceive a sequence such as bmif as one syllable and nhif as two syllables (see figure 6.4); that is, falling sonority in the onset (a violation of the SSP) will facilitate the perception of a schwa and hence a second syllable. The data for Korean speakers are shown in figure 6.4, reprinted from Berent’s figure 8.6 (2013). Similar results exist for English, but might be explained by language experience (i.e., onset clusters exist in
English and statistically English speakers hear more *bnif* than *nbif* forms; however, a statistical interpretation cannot be the explanation for the results in Korean because speakers have no experience with onset clusters. In section 6.3.1.2, we will see that speakers, without exposure to a sign language, naturally associate movements to syllable peaks, just as signers do.

6.3.1.2. Energy and Syllable Weight in Spoken Languages

Energy expenditure is also relevant for syllable weight. Syllable weight is correlated with the amount of energy expended, particularly if we examine tone languages that treat high tone as moraic. Gordon (2006) argued that syllable weight in such languages is not associated exclusively with length, but also with tonal properties that co-occur simultaneously with the vowel. In several tone languages—Crow (Kaschube 1967), Ijo (Williamson 1989), Kongo (Reh 1983), Iraqw (Mous 2007),
and Logoore (Goldsmith 1992)—a vowel plus high tone is treated as heavy. Vowels uttered with high tones have more energy than those with a low tone do. In these languages, syllables that would otherwise be treated as light, and not stress-attracting, are allowed to be stressed if they occur with a high tone. Discussions on this topic with Goldsmith about Logoore in the late 1980s led to an analysis of simultaneous syllable weight in ASL (Brentari 1990, 1998) that predates Gordon’s work on spoken languages.

6.3.2. Sonority and Syllable Weight in Sign Languages

Movements are the most visually salient part of a sign, just as vowels are the most acoustically salient part of a spoken word. Movement also plays a central organizing role at the phonological level in infants acquiring ASL from their deaf parents, and at the same as the stage of development as syllabic babbling in infants acquiring a spoken language (Pettito and Marentette 1991). Movement is used to count syllables in sign languages. The criteria for counting syllables in sign languages are outlined in (2).

(2) Syllable Counting Criteria (Brentari 1998):

a. The number of syllables in a sequence of signs equals the number of sequential movements in that string.

b. When several shorter (e.g., secondary) movements co-occur with a single (e.g., path) movement of longer duration, the longer movement is the one to which the syllable refers.

c. When two or more movements occur at exactly the same time, it counts as one syllable, e.g., ASL INFORM is one syllable containing an aperture change and a path movement.

These syllable counting criteria incorporate both issues relevant in this section. First, considering the three parameters of handshape, movement, and location syllables are counted according to the number of movements—the ordered elements aperture and setting—and not the unordered autosegmental tiers—selected fingers and major POA. Second, if two types of movements co-occur (e.g., a smaller one embedded in a single larger one), the energy is calculated across both, such that two simultaneous movements have higher energy than one. In other words, syllable count and syllable weight are based on sequential and simultaneous information.
6.3.2.1. Sonority Is Associated with Movements

The association of vowels with movements was originally proposed by Liddel (1984) because, functionally speaking, movements behave like vowels insofar as they are the medium of the signal. Without vowels, the speech signal would be of very low intensity, just as without movements, the sign signal would be low intensity (Brentari 2002). The idea that movement conveys the notion of sonority has been shown experimentally as well (Berent, Dupuis, and Brentari 2013) for both signers and nonsigners. The design of these studies was very similar to the ones reported earlier for Korean. Korean speakers were asked to judge the number of syllables, and they employed the sonority hierarchy instinctively even though their language has no CC clusters. Signers and nonsigners were asked to determine whether movements or handshapes are the default syllable nucleus. Novel signs were constructed with one or two different handshapes and with one or two different movements as shown in figure 6.5.

Analogous English examples of the ASL stimuli are given in each cell. (These were not part of the experiment and are intended to be illustrative examples of the analogous syllabic and morphological structure in a familiar spoken language.) Each experiment consisted of two tasks. In one task participants were asked to judge the number of parts without meaning (i.e., syllables) and in another to judge the number of parts with meaning (i.e., morphemes). Judgments of parts without meaning were taken to be syllabic units, and judgments of parts with meaning were interpreted as morphemes. Those in the top right and bottom left cells in figure 6.5 (outlined with bold lines) with incongruent meaningless and meaningful parts were the most informative (i.e., syllables or morphemes), because these items clearly would show differences in the participants responses if they are making different associations for handshape and movement.

The results of the four experiments are shown in figure 6.6. With just a few training examples prior to the experiment and no feedback, ASL signers readily associated movement with meaningless units (syllables) and handshape with meaningful ones (morphemes); this is evident from the interaction seen in (figure 6.6a). Nonsigners did not perform like the signers when given the same training without instruction (figure 6.6b); they seem to associate movement with both meaningless parts (syllables) and meaningful parts (morphemes). In prior work (Brentari 2006; Brentari et al. 2011) it has been shown that nonsigners tend to ignore handshape in judging the number of words in a string and to focus heavily...
on movement, so this result was not surprising. Two additional experiments with nonsigners were conducted to see whether they could be taught to pay attention to both dimensions independently (handshape and movement). One of these new studies provided nonsigners “natural” feedback after the training items; it was called “natural” because it corresponded with the judgments of the signers. Participants were instructed to associate movement with meaningless units (syllables) and handshape with meaningful ones (morphemes). With natural feedback, nonsigners were able to learn the task and reliably associate movement with meaningless units (syllables) and handshape with meaningful ones (morphemes). When provided unnatural feedback, they were not able to learn the rule (figure 6.6d). They were able to associate movement with meaningful units as instructed (note the blue line), but they did not learn to associate handshapes with meaningless units.
6.6 Responses of “one part” for novel ASL signs containing one handshape, two movements (left in each frame) and two handshapes, one movement (right in each frame) for (a) deaf ASL signers, (b) nonsigners without training, (c) nonsigners with “natural” training (training aligned with the signers responses), and (d) nonsigners with “unnatural” training (training not aligned with the signers responses). (From Berent et al. 2013.)
These results show that even without the relevant language experience (Korean speakers who have not heard CC clusters for spoken languages, and nonsigners who have never experienced sign languages), sonority is readily applied to the perceptually salient, high-energy elements of form and syllable nuclei are inferred correctly. This work also shows that despite the differences between articulators in signed and spoken language, comparative work such as this can ultimately lead to a better understanding of the universality of some of the principles of language organization, such as sonority.

Sonority hierarchies have been proposed for sign languages and many have several features in common which will be described will be described in this section (Corina 1990; Perlmutter 1992; Brentari 1993; Sandler 1993; van der Hulst 1993) (see figure 6.7). Assuming that movements made with larger articulators are produced with more energy, we see that the amount of energy correlates with proposed sonority hierarchies in sign languages: path movements produced by the elbow or shoulder are higher in sonority than are local movements, produced by the forearm, wrist, or hand. Lack of any movement (stasis) is lowest on the hierarchy, but there are no well-formed signs without a movement.

A sonority hierarchy is needed in sign languages to account for at least two phenomena. First, the sonority hierarchy is needed to explain why some ASL compounds are produced as disyllabic words and others as monosyllabic words. In ASL compounds if the first stem of the compound
has a path movement (relatively high on the sonority hierarchy), the movement is retained and the resulting compound is disyllabic. If, instead, the first stem of the compound has a trilled movement (relatively low on the sonority hierarchy), the movement is deleted and the resulting compound is monosyllabic. This can be seen in the difference between BLACK^NAME (bad reputation) versus SPEAK^NAME (mention). As a stem, BLACK is articulated with a path movement, which is retained in the compound, and the resulting compound is disyllabic. As a stem, SPEAK is articulated with a trilled movement, which is not retained in the compound, and the resulting compound is monosyllabic. Another phenomenon in ASL that requires a sonority hierarchy is the nativization of fingerspelled sequences (e.g., #P-H-O-N-O-L-O-G-Y becomes [P-H]σ [G-Y]σ, with the other handshapes deleted) (figure 6.8). In such forms, the syllable nuclei are constructed on the transitional movements between handshapes. While these transitions typically involve only finger or hand movements, a few transitional movements between letters involve the wrist and forearm. These are preferred as syllable peaks over movements of the fingers, precisely because elbow and wrist movements are larger than finger or hand movements (see figure 6.8).

Despite the fact that a sonority hierarchy is needed, there is no evidence for the SSP at work in sign languages. In the hold-movement model of sign language phonology, movements have conceptually been associated with vowels, and holds could be considered syllable margins. Perlmutter (1992) presented a proposal that used positions instead of holds, and he also observed that complex onsets or codas clusters do not exist in ASL. Due to the lack of complex clusters, coupled with the fact that syllable margins in the majority of monomorphemic cases are largely redundant (most material in the “onset” and “coda” would be the same), Meier (2002) argued, that sign languages do not employ the SSP, because, unlike spoken languages that have one major oscillator (i.e., the mandible), sign languages have many (i.e., body, arms, hand, head). This demotes the SSP in sign languages; however, sonority is one of the truly universal organizing principles of phonology.
6.3.2.2. Syllable Weight and Movement

The possibility of combining simultaneous properties for syllable weight in spoken languages was not widely known until Gordon (2006), which was the first large typological study of these phenomena. It is clear now, however, that the salience of a syllable nucleus can be determined sequentially or simultaneously in spoken languages. With regard to syllable weight with sonority in sign languages, there is a heavy/light sensitivity that is important for the application of reduplication and demonstrates that sign syllables are also sensitive to simultaneous layering of movement components as syllable weight. This distinction can be traced by to expenditure of energy if we assume that one movement element requires less expenditure of movement than two. To be concrete, a group of ASL verbs containing one movement element are permitted to undergo reduplication, resulting in derived nominal forms (e.g., the path movement in ASL SIT can be repeated to derive the nominal form CHAIR (Supalla and Newport 1978), in contrast, signs consisting of two or more movements, such as ASL INFORM, which contains both a path and aperture movement, cannot be reduplicated to obtain a derived nominal form (*INFORMATION via reduplication). This suggests that forms allowing reduplication have one simultaneous movement component and are light syllables, while those that disallow reduplication have two or more simultaneous movement components and can be therefore considered heavy. This distinction has been argued for in Finnish Sign Language as well (Jantunen and Takkinen, 2010). Light syllables readily allow reduplication, while heavy syllables do not.

Returning to the dependency model (figure 6.3c) and the prosodic model (figures 6.3c and d) of sign language phonology, in the dependency model, the root node is a single “segment” structure linked to handshape and POA (NB, segment is not equal to timing units in this model). Within that model, movement is given a minor role as a manner feature, since van der Hulst (1993) argued that movement could be derived from transitions between handshape and POA features, at least for signs that are monomorphemic. In contrast, the prosodic model ascribes a central role in sign language phonology to movements, which not only unites all features that have ordered elements, as described in section 6.2.2, but also provides a coherent backbone on which to build syllable nuclei, to calculate sonority and syllable weight, and to provide the foundation for higher-order prosodic structure.

The ASL sonority hierarchy is built into the representation of the prosodic model (figure 6.3d), with components of movements that create a...
setting change or a particular direction of movement (path movements) located higher in the representation than those with a change produced at the wrist or hand (local movements) because path movements are in general larger, more visually salient, and require more energy than those of the wrist and fingers. Further, each additional movement component is argued to be a weight unit, a mora. The relationship of the syllable, sonority, and syllable weight is organized in a transparent way in this model. This model was inspired by Goldsmith’s version of the autosegmental phonology and of the syllable.

Carstairs-McCarthy (2001) asked whether the syllable, as a unit of phonological description, is modality-neutral (with fundamentally the same sense in descriptions of signed and spoken languages), or whether the syllables of signed and spoken languages are really different phenomena. He asserted that it was legitimate to appeal to aspects of spoken syllables that are undoubtedly modality-dependent, such as their physiological underpinnings, perhaps in the number and type of oscillators available in the vocal apparatus versus the sign apparatus. Such differences are reflected in the presence versus absence in spoken and signed languages, respectively, of the SSP. By examining more basic notions, however, such as energy and perceptual salience, we see that the notion of sonority is truly universal.

6.4. Conclusion

It is rare in the history of science that we can speak of a paradigm shift, but I believe that autosegmental phonology constituted exactly that with respect to phonological representation. The constructs introduced (autosegmental tiers) or reintroduced (the syllable) by autosegmental theory have changed the way that phonology is practiced today, so much so that these ideas are often taken for granted. From the point of view of sign language phonology, autosegmental phonology made it possible to see cross-modal similarities and differences and to make predictions about structure that were not possible previously, changing theories of phonological representation (spoken and signed) ever since.

Notes

1. Goldsmith’s recent, ten-dimensional visualizations displaying eigenvector analyses of the lexical and morphological structure of English and
French are also evidence of his longstanding interest in the multidimensional space within which a grammar operates (Goldsmith 2014; Lee and Goldsmith, 2017).

2. Both the “rolladeck” structure of features as in (1a) and the hierarchical feature geometry as in (1b) are often represented as if they were flat in an attempt to make printing and viewing a bit easier on the page, but, as we know, printable simplicity is completely irrelevant to simplicity of phonological representations in the mind.

3. In the case of an aperture change, specific order is respected (⟨⟩ > ⟨⟩ not ⟨⟩ > ⟨⟩).

4. For ease of exposition I have simplified many details of both models. For example, in both models the handshape node is part of a larger structure known as the articulator, and orientation has been omitted.

References


