The Logical Nature of Phonology Across Speech and Sign

Jonathan Rawski, Stony Brook University

Abstract

This article examines whether the computational properties of phonology hold across spoken and signed languages. Model-theoretic representations of spoken and signed words, as well as logical mappings over these structures, are introduced as a powerful framework for structural and computational comparisons. Several phonological processes in sign are shown to require the same logical complexity as their spoken counterparts, suggesting an amodal sensitivity to notions of locality and memory, as well as a computational tradeoff between sequentiality and simultaneity in specific modalities. These analyses provide a necessary and sufficient condition for amodal aspects of phonology, and allow for promising new means to analyze issues of linguistic modality and the cognitive status of phonological knowledge.

1 Introduction

The presence of sign language phonology poses a challenge for theories of the phonological module, since the structure and organization of phonological form is influenced in part by the physiology of the systems that produce and perceive them. However, striking similarities have been found in phonological systems across modalities, leading some to argue for an “algebraic" phonology of computational rules (Berent, 2013). Others argue for an abstract “substance-free phonology" in either modality, which is cognitively independent from the physiology but mediates between grammatical form and perceptual form (Reiss, 2018). Sandler and Lillo-Martin (2006) take up this challenge and define phonology across modalities as “the level of linguistic structure that organizes the medium through which language is transmitted".

Sign languages arise spontaneously in deaf communities, are acquired during childhood through normal exposure without instruction, and exhibit all of the facets and complexity found in spoken languages (see Sandler and Lillo-Martin (2006) for a groundbreaking overview). However, if humans evolved for language in general without respect to modality, we should find hearing communities that just happen to use sign language instead of spoken language, and we do not. Sign languages are thus “an adaptation of existing physical and cognitive systems for the purpose of communication among people for whom the auditory channel is not available" (Sandler, 1993b).

Sign languages offer, as Sandler (1993b) puts it, “a unique natural laboratory for testing theories of linguistic universals and of cognitive organization." They give insight into the concrete contents of grammatical form, and conditions on which aspects of grammar are amodal and which are tied to the modality. Schlenker (2018), speaking about semantics, notes that “investigating Universal Semantics from the standpoint of sign language might help reconsider foundational questions about the logical core of language, and its expressive power".

How then can we study in a principled and meaningful way the expressivity of **phonology** across speech and sign? Computational characterizations of phonological processes distill the necessary and sufficient conditions of any system that performs them. Heinz (2018) writes that understanding the computational nature of phonology requires determining the nature of 1) the abstract, underlying representations, 2) the concrete, surface representations, and 3) the transformation from underlying forms to surface forms. Each clarifies the properties to which a cognitive mechanism needs to be sensitive in order to correctly classify and process forms. As an alternative to developing a specific computational model, one can determine abstract measures of complexity that "are invariant across all possible cognitive mechanisms and depend only on properties that are necessarily common to all computational models that can distinguish a pattern" (Rogers et al., 2013)

This article analyzes the nature of such representations by using elements of Finite Model Theory, a branch of mathematics known in semantics but not often applied in the study of linguistic form. Model Theory refines the computational characterization by describing the **content** of linguistic structures. The computational nature of processes over these representations is analyzed with statements in mathematical logic. These are well-established methods that have been used in linguistics before to characterize and compare particular grammatical theories in syntax and phonology (Rogers, 1998; Potts and Pullum, 2002; Pullum, 2007; Graf, 2010).

Why the focus on logic? The connections between language, mathematical logic and computational models have a long research history, dating back to the foundations of theoretical computer science. Some of the earliest results in linguistic theory established the place of natural language patterns within the theory of computable functions (Chomsky, 1956). Later, it was discovered that phonological rules and constraints are entirely describable by finite-state machines (Johnson, 1972; Kaplan and Kay, 1994), meaning the computation is characterized by a bounded memory. This is important because foundational work by Büchi (1960) showed that every finite-state machine is equivalent to statements in monadic second-order logic. This logic-computation connection shows how to relate the **specification** of a system behavior (as given by a logical formula) to a possible **implementation** (as the finite-state behavior), a distinction used in the Declarative Phonology framework (Scobbie et al., 1996). One can understand the nature of a process in terms of the information and computation needed to perform it, especially for representations where the structure of the automaton is not readily apparent.

The goal of this article is to introduce the model-theoretic and logical approach to phonology, and to apply this approach to the problem of modality and the nature of phonology across modalities. The article proceeds as follows. Section 2 overviews the model-theoretic approach to phonological representation and considers various representations of spoken and signed words. Section 3 describes phonological processes as logical transformations over word models. Section 4 showcases the logical power of several widely attested phonological processes in speech and sign. Section 5 discusses the implications of these results for debates on sequentiality and simultaneity. Section 6 discusses some implications for the cognitive status of phonological knowledge.

2 Model-Theoretic Representations of Words and Signs

This section defines the central ideas of model-theoretic representations and considers various representations of words and signs. This involves deciding what kind of objects we are reasoning about and what relationships between them we are reasoning with. Model theory provides a unified ontology and a vocabulary for representing many kinds of objects, by considering them as **relational structures** (see Libkin (2004) for a thorough introduction). This allows flexible but precise definitions of the structural information in an object, by explicitly defining its parts and the relations between them. This makes model-theoretic representations a powerful tool for analyzing the information characterizing a certain structure.

For example, one common way of representing phonological words is as a string. Strings are sequences of events, each labeled with particular symbols that describe **properties** of those events. These properties can be the speech sound symbols in the IPA, phonological features, or a set of orthographic symbols, but for the present purposes let us, without loss of generality, consider a set of properties $\Sigma = \{a, b, c\}$. Strings are combinations of these symbols at certain events, like the word *baba*.

Model-theoretic representations for finitely-sized objects like strings contain two parts. The first is a finite set of elements $\mathcal{D}$, called the **domain**, taken here to be elements of the natural numbers, as is common. The second is a finite set of k-ary **relations** $\mathcal{R}$, and **functions**[1] $\mathcal{F}$, which are subsets of the domain. The relations and functions provide information about the domain elements. The **model signature** $\mathcal{M} = \langle \mathcal{D}; \mathcal{R}; \mathcal{F} \rangle$ collects these parts and defines the nature of the model in terms of the information in the representation. One model signature for words, called the **successor model**, is given below in (1)

(1) $$\mathcal{M}^S \stackrel{\text{def}}{=} \big\langle \mathcal{D}; \mathcal{R} = \{R_\sigma \mid \sigma \in \Sigma\}; \mathcal{F} = \{\mathtt{p}(x), \mathtt{s}(x)\}\big\rangle$$

This model says that for every property σ in the set of properties Σ, there is a unary relation R_σ in $\mathcal{R}$ that can be thought of as a labeling relation for that symbol. For our set $\Sigma = \{\mathtt{a}, \mathtt{b}, \mathtt{c}\}$, $\mathcal{R}$ includes the unary relations R_a, R_b, R_c. The two unary functions in $\mathcal{F}$, $\mathtt{p}(x)$ and $\mathtt{s}(x)$, describe a linear order over the domain elements by picking out the immediate **predecessor** and **successor** of some given position, respectively. In general, $\mathtt{s}(x) = x + 1$ and $\mathtt{p}(x) = x - 1$. The predecessor function is a total function, because it is defined so that the initial position is its own predecessor, i.e. $\mathtt{p}(0) = 0$. Similarly, the final position is its own successor, so the successor function is also total. Then in a string of n positions, $\mathtt{s}(n) = n$.

As an example, the model for the word $baba$ under this theory is denoted $\mathcal{M}^S_{baba}$. For the properties $\Sigma = \{\mathtt{a}, \mathtt{b}, \mathtt{c}\}$, $\mathcal{M}^S_{baba}$ is defined and represented visually in Figure 1. Here the word model's domain $\mathcal{D}$ consists of four nodes, each as a node with an index below it. Unary relations are illustrated as node labels. For example, node 2 is labeled a. The successor and predecessor functions are illustrated by directed edges (arrows), with the s or p label, respectively (i.e. s(1) = 2).

Model vocabulary items are also called **atomic formulas**, because they are the primitive terms from which larger logical expressions are built. Let x, y, etc., be variables and then let $\mathtt{p}(x) = y$, $\mathtt{s}(x) = y$ and $R_\sigma(x)$ for each σ in the set of properties be atomic predicates. These variables will be assigned values in $\mathcal{D}$ in a model. $\mathtt{s}(x) = y$ will evaluate to true if and only y is the successor of x in $\mathtt{s}(x)$ in that model. For example, $\mathtt{s}(x) = y$ is true in Fig. 1 when x is assigned to 1 and y is assigned to 2, but not when x is assigned to 1 and y is assigned to 4. Similarly, $\mathtt{a}(x)$ is true in a

[1] It is true that every k-ary function can be considered as a $(k+1)$-ary relation, and every constant as a 0-ary function or (singleton) unary relation, but for explicitness I keep them this way

$$
\begin{aligned}
&\mathcal{M}^{S}_{baba}\\
\mathcal{D} &= \{1,2,3,4\}\\
R_a &= \{2,4\}\\
R_b &= \{1,3\}\\
R_c &= \{\emptyset\}\\
\mathtt{s}(x) &= \begin{cases} 2 & x=1\\ 3 & x=2\\ 4 & x\in\{3,4\}\end{cases}\\
\mathtt{p}(x) &= \begin{cases} 1 & x\in\{1,2\}\\ 2 & x=3\\ 3 & x=4\end{cases}
\end{aligned}
$$

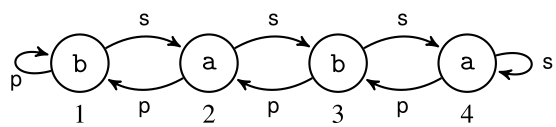


Figure 1: Successor word model for *baba*

model only when x is assigned to an element in $\mathtt{a}$ in that model. For example, $a(x)$ is true in Fig. 1 when x is assigned to 2 (denoted $2 \in R_a$, $R_a(2) = \mathtt{TRUE}$, or, equivalently, $\mathtt{a}(2) = \mathtt{TRUE}$), but not when it is assigned to 1. I also add an atomic predicate $x = y$ that evaluates to true when x and y are assigned to the same element in $\mathcal{D}$.

Atomic formulas are also predicates, so for a model theory we can implicitly define the full set of Boolean connectives such as conjunction, disjunction, implication, and negation, as well as quantifiers. This means that given a model theory $\mathcal{M}$, logical predicates can be defined to make it easier to refer to certain types of information in $\mathcal{M}$.

For example, if the atomic properties of the sequential elements are taken to be a set of phonological features like $\mathtt{voiced}(x)$ or $\mathtt{stop}(x)$, one may use multiple unary relations to describe the same domain element, and therefore refer to individual phonemes with user-defined predicates, such as in (2-3). One might also wish to pick out certain privileged elements in the word, such as word edges, by picking out **constants**, as in (4-5). Constants may also be a part of the model signature. A modified representation of the word *baba* in the successor model using these predicates is shown in Figure 2.

(2) $$\mathtt{b}(x) \stackrel{\text{def}}{=} \mathtt{voiced}(x) \wedge \mathtt{labial}(x) \wedge \mathtt{stop}(x)$$

(3) $$\mathtt{a}(x) \stackrel{\text{def}}{=} \mathtt{vowel}(x) \wedge \mathtt{back}(x) \wedge \mathtt{low}(x)$$

(4) $$\mathtt{first} \stackrel{\text{def}}{=} \mathtt{p}(x) = x$$

(5) $$\mathtt{last} \stackrel{\text{def}}{=} \mathtt{s}(x) = x$$

The model-theoretic perspective gives us a flexible position from which to compare the representational content of spoken and signed phonological words. One obvious strategy is to say that that the representation for signs is essentially equivalent to that of spoken words, i.e. they

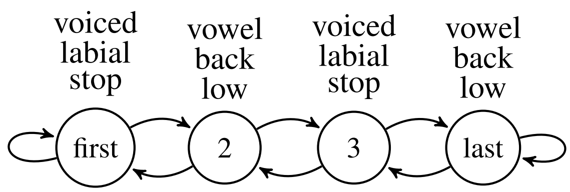
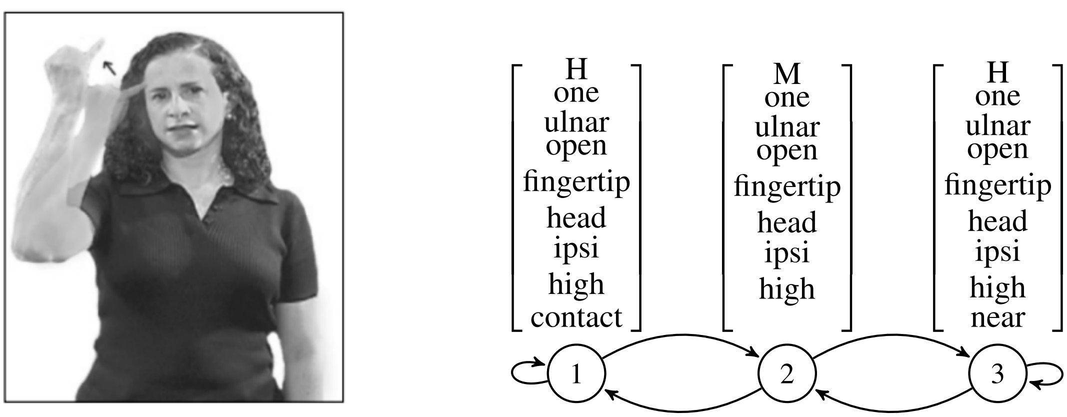


Figure 2: Visual of feature-based word model for 'baba' with word-edge constants

are strings. Virtually all models of sign phonology allow sequentiality, as a sequence of static and dynamic segments (Liddell, 1984; Sandler, 1986; Liddell and Johnson, 1989; Perlmutter, 1993), or a sequence of abstract timing units where only the non-dynamic endpoints ultimately associate (van der Hulst, 1993; Brentari, 1998).

This is essentially the position taken by Liddell (1984), who represents signs as strings of Hold and Movement segments (usually 3) labeled with various phonological features. Model-theoretically, this just means adapting our set of properties and thus the unary relations in the signature. This representation explicitly makes the claim that the only difference between spoken and signed representations is the size and content of the feature system. However, considering the string-based representation of the ASL sign IDEA given in Figure 3, it is easily apparent that such representation is highly redundant, as most featural information is stable across the whole sign.



IDEA (ASL)

Figure 3: Left: ASL 'IDEA' (Image copyright Diane Lillo-Martin and Wendy Sandler). Right: Visual of string based word model for 'IDEA'

Advances in phonological theory liberated certain feature content from a single temporal dimension by positing **Autosegmental Representations** (Williams, 1976; Goldsmith, 1976). In ARs, utterances are made up of several kinds of simultaneous levels, with each level related to but ordered independently of any other level. Phonological primitives are arranged in distinct strings or tiers, with an association relation relating units on different tiers. ARs have been argued to provide natural accounts of many tone and segmental processes because the non-local interactions between elements on a tier are always mediated through associations to a common element on another tier (see Jardine (2016b, 2017a) for a formal account).

To augment the model theory for words in order to make them ARs, Chandlee and Jardine (2019) note that all that is required is to add a binary association relation $\mathtt{A}(x, y)$ to the model signature. This results in a new model signature, $\mathcal{M}^{AR}$, distinct from the previous string model signature,

(6) $$\mathcal{M}^{AR} \stackrel{\text{def}}{=} \langle \mathcal{D}; \{\mathtt{a}, \mathtt{b}, \mathtt{H}, \mathtt{L}\}; \{\mathtt{A}(x, y)\}; \{\mathtt{p}(x), \mathtt{s}(x)\}\rangle$$

A model of our string *baba* and a visual representation are given in Figure 4. Here the number of domain elements has increased, and some are now only labeled with tonal features and some with segmental features. Successor and predecessor functions hold for all elements, and on different tiers. Note that now two elements precede themselves and two succeed themselves, though this does not have to be the case. The association relation $\mathtt{A}(x, y)$ holds between elements on the two tiers, shown using double-arrows. This captures a structure where *baba* has a high tone H associated to the first vowel, and a low tone L associated to the second.

$$\mathcal{M}^{AR}_{baba}$$
$$\mathcal{D} = \{1, 2, 3, 4, 5, 6\}$$
$$\mathtt{A}(x, y) = \{\langle 2, 5\rangle, \langle 4, 6\rangle\}$$
$$\mathtt{a} = \{2, 4\}$$
$$\mathtt{b} = \{1, 3\}$$
$$\mathtt{H} = \{5\}$$
$$\mathtt{L} = \{6\}$$
$$\mathtt{s}(x) = \begin{cases} 2 & x = 1 \\ 3 & x = 2 \\ 4 & x \in \{3, 4\} \\ 6 & x \in \{5, 6\} \end{cases}$$
$$\mathtt{p}(x) = \begin{cases} 1 & x \in \{1, 2\} \\ 2 & x = 3 \\ 3 & x = 4 \\ 5 & x \in \{5, 6\} \end{cases}$$

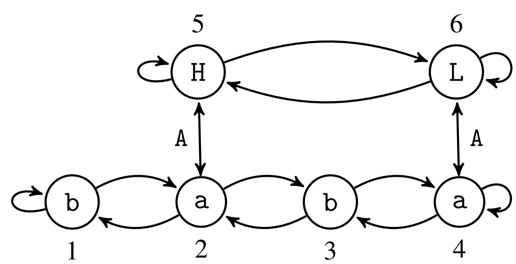


Figure 4: Autosegmental word model for “bábà"

While some sequential structure is acknowledged in almost all models of signs, sign representations have increasingly been argued to be inherently autosegmental in nature. Liddell and Johnson (1989) presented an updated version of their Move-Hold model where handshape features are autosegmentally associated to Movement and Hold elements on a segmental tier. Sandler (1986, 1989), demonstrated that hand configurations can be independent morphemes (classifiers), and exhibit autosegmental stability phonologically. She proposed the Hand Tier model, which represents sequential Location (L) and Movement (M) segments on one tier, allowing explicit reference to sequential information, and represents Handshape Configuration (H) autosegmentally, where one handshape characterizes the whole sign, in contrast to the one-per-segment in the Move-Hold model. ARs for place of articulation features are a more recent innovation, seen in Brentari's (1998) Prosodic Model and in van der Hulst's Dependency Phonology Model (van der Hulst, 1993, 1994).

Incorporating these autosegmental attributes of sign into a model theory is again straightforward, since it simply amounts to adding more relations or functions. An AR model of the sign, $\mathcal{M}^{SAR}$, is shown in 7.

(7) $$\mathcal{M}^{SAR} \stackrel{\text{def}}{=} \langle \mathcal{D}; \{\mathtt{L}, \mathtt{M}, \mathtt{H_i}, \mathtt{P_j}\}; \{\mathtt{A}(x,y)\}; \{\mathtt{p}(x), \mathtt{s}(x), \mathtt{loc}(x)\} \rangle$$

I include Location, Movement, Handshape, and Place as unary relations $\{\mathtt{L}, \mathtt{M}, \mathtt{H_i}, \mathtt{P_j}\}$, where the subscripts i, j mean they are one out of a set of possible handshapes or places, respectively. The association relation $\mathtt{A}(x, y)$ associates elements on the LM-tier to the handshape tier, and the function $\mathtt{loc}(x)$ relates L segments to their specific segment on the Place tier. The use of a function here is not necessary; one could easily accomplish the same thing with another binary relation, but I use a function to demonstrate the method. A representation of a monosyllabic sign with one place feature and one handshape feature is given in Figure 5

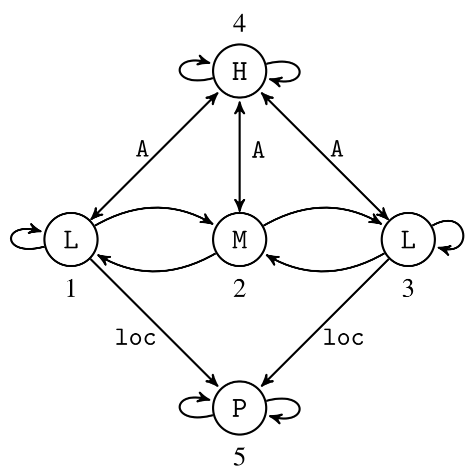


Figure 5: Autosegmental word model of a monosyllabic sign

In what follows I will utilize the autosegmental model $\mathcal{M}^{SAR}$ for signs. I am, however, not claiming that this is the correct model of spoken or signed phonological words, or that this is supposed to be a perfect formulation of one of the aforementioned theories — far from it. The proper characterization of the phonological word is very much an open question, even for spoken phonology.

What I am saying is that Model Theory allows for freedom and preciseness in choosing the representations (embedded in the atomic predicates) that one takes to be linguistically relevant, and in the strong case, cognitively real, and to see what effect these have on the structures that are allowable. If one has a commitment to a certain phonological representation, say graphs for feature geometries, or trees for representing syllabic or prosodic constituencies, one can represent them model-theoretically and examine their effects on the computation. The earlier mentioned logic-automaton connection for finite string models holds for many different data structures, including infinite strings, trees, infinite trees, hypergraphs, graphs of bounded tree-width, traces (a syntactic model for concurrency), texts (strings with an additional ordering), among others (see Thomas (1997) for an overview of formal languages within the framework of mathematical logic).

3 Logical Mappings and Computational Expressivity

Model-theoretic representations provide a clear path to define the nature of transformations from underlying to surface forms. The theory of logical transformations states that we can define a mapping from the set of input structures in one model signature to the set of output structures in another model signature by defining each relation in the signature of the output in terms of the logic of the input (Courcelle, 1994; Engelfriet and Hoogeboom, 2001; Filiot and Reynier, 2016). For an input signature $\mathcal{M}^I$ and an output signature $\mathcal{M}^O$, a logical mapping $\mathcal{T}$ specifies a finite number of output **copies** of the domain $\mathcal{D}$, and for each copy defines each function, relation, and constant in $\mathcal{M}^O$ in terms of the input $\mathcal{M}^I$. Mappings additionally define a **licensing function** that specifies which domain elements survive in the output. If there are multiple copies of the input domain, one must specify relations between them. Since each of these formulas are terms, they are semantically interpreted with respect to the input structure.

Consider a mapping that changes $\texttt{b}$'s in a word model to $\texttt{a}$'s. Here $\mathcal{M}^I = \mathcal{M}^O = \mathcal{M}^S$, the successor model. Then given the set of properties $\Sigma = \{\texttt{a}, \texttt{b}, \texttt{c}\}$, this mapping $\mathcal{T}_{ba}$ defines the set of predicates over the input structure where the superscript O indicates the relations over the output. (8) says that a domain element is labeled an $\texttt{a}$ in the output iff it was labeled an $\texttt{a}$ or $\texttt{b}$ in the input. (9) specifies which output elements have the $\texttt{b}$ relation, and is $\texttt{FALSE}$, since no outputs satisfy it. (10-11) say that the predecessor and successor functions are true for any input-output pair as they were in the input. For example, using the model from above, when applied to the input $\mathcal{M}^S_{baba}$, the mapping changes the label of the first and third positions from $\texttt{b}$ to $\texttt{a}$ and leaves the remaining positions unchanged, as shown below.

(8) $$\texttt{a}^O(x) \overset{\text{def}}{=} \texttt{a}(x) \vee \texttt{b}(x)$$

(9) $$\texttt{b}^O(x) \overset{\text{def}}{=} \texttt{FALSE}$$

(10) $$\texttt{p}^O(x) \overset{\text{def}}{=} \texttt{p}(x)$$

(11) $$\texttt{s}^O(x) \overset{\text{def}}{=} \texttt{s}(x)$$

(12) $$lic^O(x) \overset{\text{def}}{=} \texttt{TRUE}$$

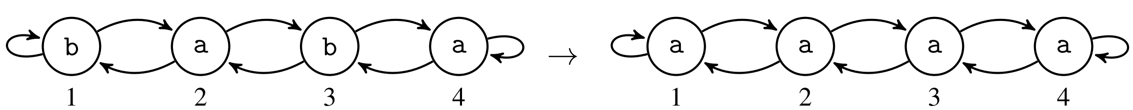


Figure 6: Visual of input and output word models for the mapping in (4-8)

Notice that none of the statements in the preceding mapping involve the use of quantifiers. Such statements are quantifier-free fragments of first-order logic, and a mapping that is completely quantifier-free is called a **Quantifier-Free mapping** (QF). Other types of logic may be used to similar effect. Statements in First-Order (FO) logic can add universal and existential quantifiers to quantify over elements of the domain. Monadic Second-Order (MSO) statements can also quantify over *sets* of domain elements[2]. Importantly, these are logical statements about models, not about the logical axioms that define parts of the model signature. The logical power of these axioms, and the limitations that poses on the models or computations defined is an important open question for phonology, one uniquely brought out by the model-theoretic framework.

There is a direct link between the type of logic used in a mapping and the computational power required of that mapping. It was stated earlier that MSO statements over model signatures are equivalent to Finite-State acceptors, which define the Regular languages (Büchi, 1960). The set of MSO-definable mappings uniquely defines 2-way finite-state transducers (machines with input and output symbols), which compute the set of **regular relations** (Engelfriet and Hoogeboom, 2001). There is thus a deep connection between MSO logic and the property "being regular", meaning that the amount of memory required does not grow with the size of the input. Weaker logics define subclasses of the regular languages and relations (see Rogers and Pullum (2011) for an overview of subregular languages, and Filiot (2015) for transducers).

To see why quantification is important, compare (13) to (14), adapted from Strother-Garcia (2018). The former states that an output position x will be labeled $\mathtt{a}$ if the corresponding input position is an $\mathtt{a}$ *or* if there is a position labeled $\mathtt{b}$ *somewhere* in the input. Checking whether $\mathtt{a}^O(x)$ is true requires evaluation of the entire string to see if any position is labeled $\mathtt{b}$. This is due to the existential quantifier $\exists$, which makes (13) strictly FO. In contrast, (14) lacks any quantification. $\mathtt{a}^O(x)$ can be evaluated independently at every position in the string.

(13) $$\mathtt{a}^O(x) \stackrel{\text{def}}{=} \mathtt{a}(x) \vee (\exists y)[\mathtt{b}(y)]$$

(14) $$\mathtt{a}^O(x) \stackrel{\text{def}}{=} \mathtt{a}(x) \vee \mathtt{b}(x)$$

This example illustrates the connection between quantification and locality. Computing the truth value of a predicate that involves quantification requires **global** evaluation. If the predicate does not use quantifiers, however, its truth evaluation must be possible over a substring of bounded size. Thus a QF mapping amounts to a constraint-checking function that operates **locally**, within a bounded window of evaluation. Defining a QF formula that identifies targets of the process in terms of information present in the input provides a rigorous, independent notion of what it means for a process to be local.

Chandlee and Lindell (2016) showed that the set of mappings satisfiable by QF formulas over the successor model $\mathcal{M}^S$ share a deep connection to a proper subset of the regular relations, the Input Strictly Local (ISL) functions. ISL functions determine an output string for an input based only on contiguous substrings of bounded length (Chandlee and Heinz, 2018). For example, intervocalic voicing is ISL since it only has to be sensitive to 3-segment input substrings of the form VTV. Phonologically, this class is extremely relevant, because Chandlee (2014) showed that a full 95% of the process in P-base (Mielke, 2007), a comprehensive database of phonological processes, are

[2] For formal definitions of MSO and FO, see Enderton (2001), Fagin et al. (1995), and Shoenfield (1967).

ISL functions. Additionally, ISL functions have efficient learning algorithms from positive data (Chandlee et al., 2014) and are learned more easily by humans in learning experiments (Finley, 2009). Intuitively, the QF-ISL connection stems from the fact that all the needed information can be found within a bounded window of material surrounding a given input position. Just as in the MSO-Regular case, there is a deep connection between Quantifier-Freeness and Strict Locality, because both rigorously define the notion of “being local".

As an example, Strother-Garcia (2018) provides an analysis of syllabification in Imdlawn-Tashlhiyt Berber, and shows that it can be succinctly captured using a QF mapping. This is intriguing, since Berber syllabification traditionally was a strong motivation for global optimization in constraint-based frameworks, but the process at heart turns out to be a local .

Chandlee and Jardine (2019) extend QF mappings to consider Autosegmental model signatures. They show that an input-output map is Autosegmental-Input Strictly Local (A-ISL) if it can be described with a QF mapping where the model signature is an AR. Model-theoretically, this generalizes the notion of locality from considering “substrings" to “**sub-structures**", where now the chunks that are being evaluated are members of an autosegmental graph. The A-ISL class is more powerful than the ISL class, but each preserves the notion of Strict Locality. They further prove that if an AR map is A-ISL, then the individual map on each tier is an ISL function.

The flexibility given by the model-theoretic perspective for defining linguistic representations, combined with the precise connections between logical statements and computation, enables a powerful ability to characterize the nature of phonological (and indeed any linguistic) processes across modalities. Model-theoretically, representations can be defined for each modality on its own terms, using the information characteristic of that modality. Parallels and divergences then emerge in a way that enables precise comparison. One can then understand the nature of a process in terms of the minimal information and computation needed to perform it.

4 Logical Mappings in Sign Language Phonology

This section provides logical characterizations of several phonological processes in sign language, in order to compare their complexity to their spoken equivalents. Two of the processes (metathesis and partial reduplication) were analyzed by Chandlee (2014) in spoken language and shown to be Input Strictly Local functions. The third, compound reduction, is similar to blending in speech, and relies extensively on the simultaneous character of the handshape configuration and place features. Rawski (2017) showed that all of these mappings are Input Strictly Local when signs are represented as strings. The following analyses enrich the structure to encompass the autosegmental character of the sign. Examples are drawn from American Sign language (ASL), but these processes are widely attested across sign languages. In the definitions that follow, I will often omit formulas for relations and functions whose definition is `TRUE` or identical to their input formula.

4.1 Location Metathesis

One sign process making crucial reference to discrete locations is metathesis, which switches the first and last locations of a sign (Liddell and Johnson, 1989). In signs in which the signing hand makes contact at two different settings within one major body area (such as the head or chest), the order of the two may be reversed. Liddell and Johnson claim that the conditioning environment for such metathesis is the location of the preceding sign. For example, the ASL sign ‘DEAF’ starts at the upper cheek and moves down to contact the chin, in citation form and in most contexts, as shown in Figure 7a following ‘FATHER’, which is signed at the forehead. In 7b, ‘DEAF’ follows ‘MOTHER’, a sign made at the chin, and in this context ‘DEAF’ begins with the chin location, and ends at the cheek.

a. FATHER DEAF (ASL) b. MOTHER DEAF (ASL)

Figure 7: Contact Metathesis in ASL ’DEAF’ (Images copyright Diane Lillo-Martin and Wendy Sandler)

Because of the intervening Movement segment within the metathesized syllable, this is a particular type of non-local metathesis where the intervening segment is bounded. This type of metathesis is widely present in spoken languages. Chandlee (2014) cites an example from Cuzco Quechua Davidson (1977), where a liquid and glide segment exchange positions over an intervening vowel:

(1) ruyaq $\rightarrow$ yuraq ‘white’

Chandlee showed that such metathesis processes are Input Strictly Local, provided that the distance between segments that metathesize is bounded. This condition appears to be met in all synchronic cases.

At face value, this mapping is rather basic. I adopt a version of the sign model signature $\mathcal{M}^{SAR}$ spelled out in figure 5. To define the mapping, I specify a set of predicates defining the output structure in terms of the input. No structural information changes, and all domain elements are licensed, so I omit those relations for spatial reasons. Only the relations switching the two `P` segments need to be defined. The relations in 15 and 16 accomplish this by relabeling `P` segments in the output to their opposite, shown in Figure 8.

(15) $$P_1^O(x) \stackrel{\text{def}}{=} P_2(x)$$

(16) $$P_2^O(x) \stackrel{\text{def}}{=} P_1(x)$$

While a mapping like this captures the flavor of the metathesis process, it does not capture the notion that the conditioning environment is the location of the preceding sign, nor that the location

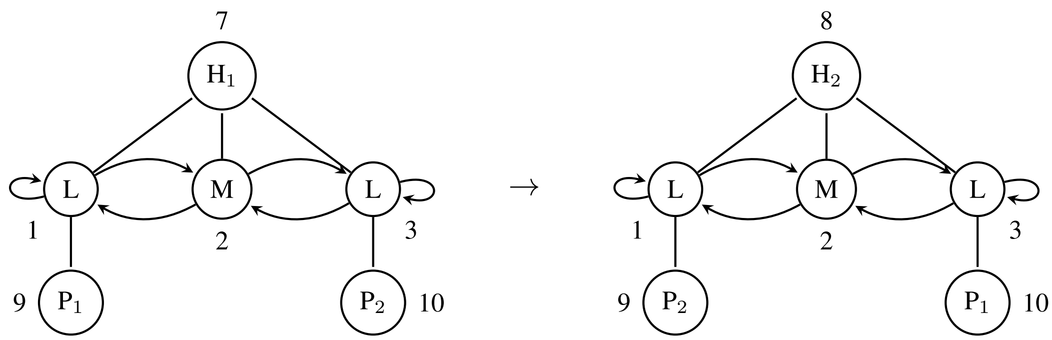


Figure 8: Input and output models for a simplified metathesis relation

change is structural, not just a change in node label. We can incorporate this information into the input structure, shown graphically in Figure 9. The model signature is equivalent to that defined in Figure 5, but here the specific model of ASL 'MOTHER DEAF' describes two signs with a movement segment in between them. Here the place features are a many-to-one function. These representational choices are supported by Liddell and Johnson (1989), who note that connected signs have an epenthesized movement segment between them prior to metathesis, and Sandler (1993a) notes that metathesis happens after "linearization" of the sign, taken to be tier conflation, where segments on separate tiers become part of the same tier.

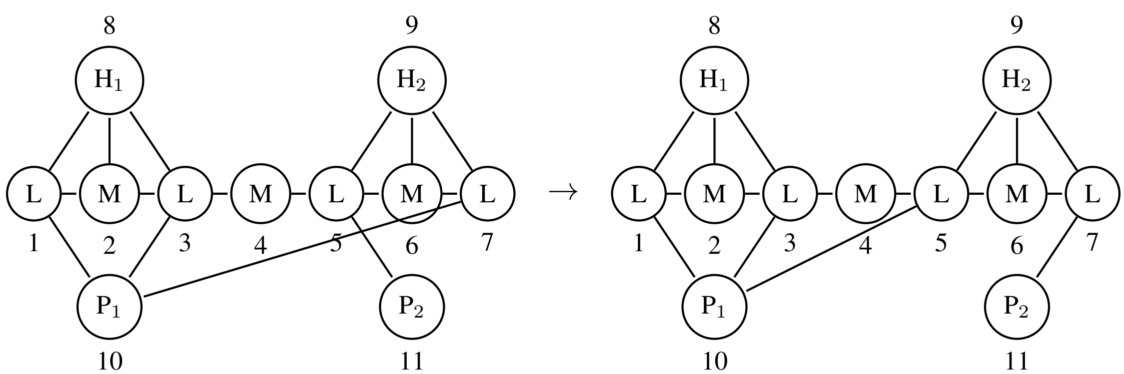


Figure 9: Visual of Input (left) and Output (right) models of the metathesis process

Now a metathesis mapping may be defined structurally. It is easy to see that almost all of the structural information stays the same. The important mapping is the location function, which is spelled out in (17). This function reassociates segments on the LM tier to elements on the place tier based on the locations of a finite amount of their successors and predecessors. In the mapping in Figure 9, the location of domain elements 5 and 7 reassociate based on the location of 3.

(17)
$$\texttt{loc}^O(x) \stackrel{\text{def}}{=} \begin{cases} \texttt{loc}(\texttt{s}(\texttt{s}(x)) & \texttt{loc}(\texttt{p}(\texttt{p}(x)) = \texttt{loc}(\texttt{s}(\texttt{s}(x))) \\ \texttt{loc}(\texttt{p}(\texttt{p}(x)) & \texttt{loc}(x) = \texttt{loc}(\texttt{p}(\texttt{p}(\texttt{p}(\texttt{p}(x))))) \\ \texttt{loc}(x) & else \end{cases}$$

This transformation ensures that if the place features of the ends of two signs agree, the second will metathesize with the intervening location segment. The input domain "size" that this mapping requires is bounded. Since all of the logical formulas over either of these autosegmental representation are quantifier-free, the mapping is A-ISL. This combined with the earlier mentioned result that metathesis is ISL across modalities with a string-based representation, shows that the nature of metathesis is computationally the same across modalities even considering different representational choices.

4.2 Compound Reduction

Many lexicalized sign compounds undergo a type of phonological reduction to preserve the monosyllabic character of canonical signs (Frishberg, 1975). Compound reduction is an amalgam of several processes. Often, sequential segments of both members of the compound delete (Liddell, 1984; Liddell and Johnson, 1989), the hand configuration of the first member also deletes, and the hand configuration of the second member spreads to characterize the whole surface compound (Sandler 1986; 1989). Other compounds reduce in different ways. Some maintain all segments and both hand configurations. Others reduce segmental structure only, maintaining two hand configurations(though see Lepic (2015)).

As an example, consider the ASL compound 'BELIEVE' ('THINK' + 'MARRY'). This compound is characterized by regressive total handshape spreading from the second sign to the first, deletion of sequential elements, and coalescence of the signs such that place information is uniquely specified for both L segments, as shown is Figure 10.

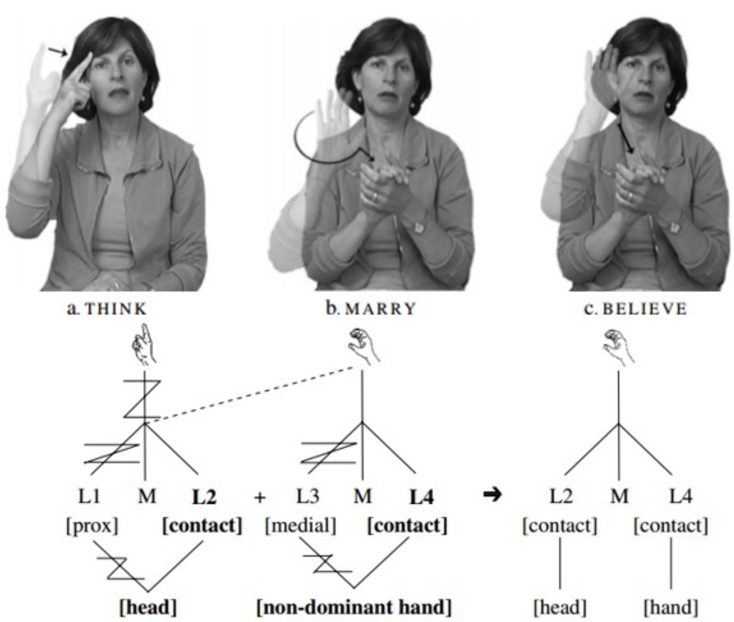


Figure 10: Compound Reduction for ASL 'THINK'+'MARRY'='BELIEVE' (TOP) and Hand Tier model representations (Bottom) (Images copyright Diane Lillo-Martin and Wendy Sandler.)

An input model theory for this compound requires a similar signature as before, so both sign elements of the compound are fully present. Since this is an underlying representation of the fully specified morphological compound, and thus no epenthetic M segment, they are represented as immediately adjacent. Importantly, I include the constants first and last in the signature as defined in (4-5), to pick out the first and last elements on the segmental tier, respectively. The complete model signature is defined in (18), and a visual representation is on the left in Fig. 11.

(18) $\mathcal{M}^{SAR} \stackrel{\text{def}}{=} \langle \mathcal{D}; \{H_i, P_i, L, M\}; \{A(x,y)\}; \{\texttt{pred}(x), \texttt{succ}(x), \texttt{loc}(x)\}; \{\texttt{first}, \texttt{last}\}\rangle$

Now we can define the reduction mapping from this input model to an output. Again, I will only define relevant transformations rather than the whole model. The first part of the mapping concerns the handshape spreading. (19) states that elements on the LM and H tiers are associated in the output if they are associated in the input to the second handshape, or if they are one of the three elements preceding a segment that was so associated. In this way, the handshape spreading proceeds locally, as it only has to account for a bounded distance between the element it is considering and one that was associated in the input.

(19)
$$\begin{aligned} A^O(x,y) \stackrel{\text{def}}{=} & [A(x,y) \wedge H_2(y)] \vee \\ & [A(\texttt{s}(x),y) \wedge H_2(y)] \vee \\ & [A(\texttt{s}(\texttt{s}(x)),y) \wedge H_2(y)] \vee \\ & [A(\texttt{s}(\texttt{s}(\texttt{s}(x))),y) \wedge H_2(y)] \end{aligned}$$

The deletion of timing and handshape segments is handled by the licensing function $lic(x)$, which specifies the domain elements that survive in the output. For the metathesis case, the licensing function always evaluated to TRUE. Here it picks out specific elements based on their properties. I specify licensing functions for each tier, and then a more general function. (20) says that if an element is a handshape, it is licensed only if it is associated to the final element of the compound. (21) says that elements labeled L or M are licensed if they are final, or the two elements which precede the final element. (22) says that a place tier element is licensed if it is the location of the last element in either of the compounding words. (23) says that in general, an element is licensed only if it satisfies one of these conditions.

(20) $$lic_H(x) \stackrel{\text{def}}{=} \bigvee_i H_i(x) \wedge A(x, \texttt{last})$$

(21) $$lic_{LM}(x) \stackrel{\text{def}}{=} (L(x) \vee M(x)) \wedge [x = \texttt{last} \vee x = \texttt{p}(\texttt{last}) \vee x = \texttt{s}(\texttt{s}(\texttt{first}))]$$

(22) $$lic_P(x) \stackrel{\text{def}}{=} \bigvee_j P_j(x) \wedge [x = \texttt{loc}(\texttt{last}) \vee x = \texttt{loc}(\texttt{s}(\texttt{s}(\texttt{first})))]$$

(23) $$lic(x) \stackrel{\text{def}}{=} lic_H(x) \vee lic_{LM}(x) \vee lic_P(x)$$

Again, note that this mapping is quantifier-free, because each formula in it is quantifier-free. There are many more types of compound reduction, involving partial handshape assimilation, and/or deleting different elements, but the input and output structure is always bounded. This ensures that the output compound structure is only determined by a finite amount of information

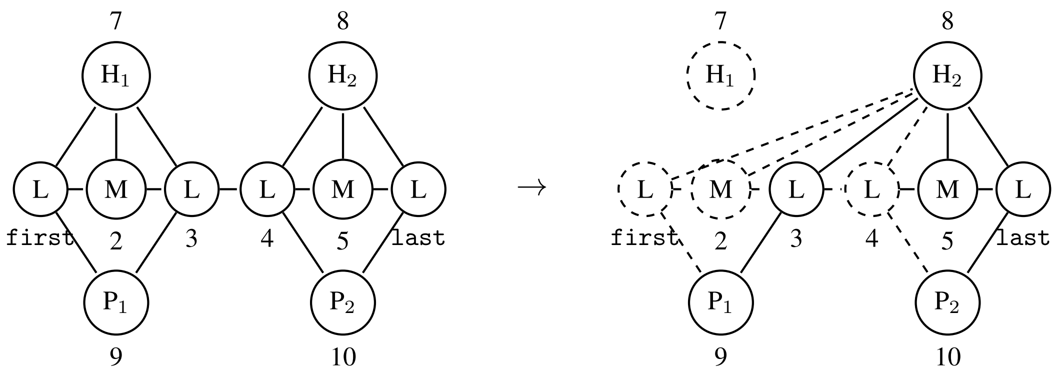


Figure 11: Visual of Input (left) and Output (right) models for compound reduction. Unlicensed elements are dashed

present in the input structure, and so does not need the power of quantification to describe the mappings. As stated in the preceding section, this means the mapping is A-ISL. Thus the notion of locality required for compound reduction is the same even when expanding the representation to include a complex interaction of processes, showcasing the local nature of the process even with additional relational structure, a point which has been made for ISL functions and opaque processes (Chandlee et al., 2018).

4.3 Final Syllable Reduplication

The nature of the output of compound reduction plays an additional role in other morphological processes. Many aspectual inflections in sign, in particular those expressing duration or iteration, involve total reduplication of the monosyllabic base (Klima and Bellugi, 1979). However, when compounds are reduplicated, the reduplicated element is the final syllable (Sandler, 1989). If the compound is reduced and monosyllabic, like ASL ‘FAINT’, then the whole form is reduplicated, shown in various forms in Fig. 12. However, if the compound is disyllabic, like ASL ‘OVERSLEEP’, only the final syllable is reduplicated. It doesn’t matter whether the last syllable has path movement only or internal movement only; each type of movement is regarded as a syllable nucleus by this reduplicatory process. The second member of the ASL compound ‘ACCIDENT’ (‘WRONG’+‘HAPPEN’) has internal movement only, specifically, orientation change. It is that syllable that gets reduplicated, and not the whole compound. In the representation of ‘ACCIDENT’, the M in parentheses is epenthetic.

Chandlee (2014) describes a class of reduplication patterns of this sort as local reduplication, since the reduplicant is affixed adjacent to the portion of the base it is copied from. Another category that meets this condition is where the reduplicant is a suffix copied from the end of the base. Chandlee cites reduplicative prefixation in Tagalog, and reduplicative suffixation in Marshallese, shown in Ex. 2 and 3, respectively, along with rewrite rules for the particular processes:

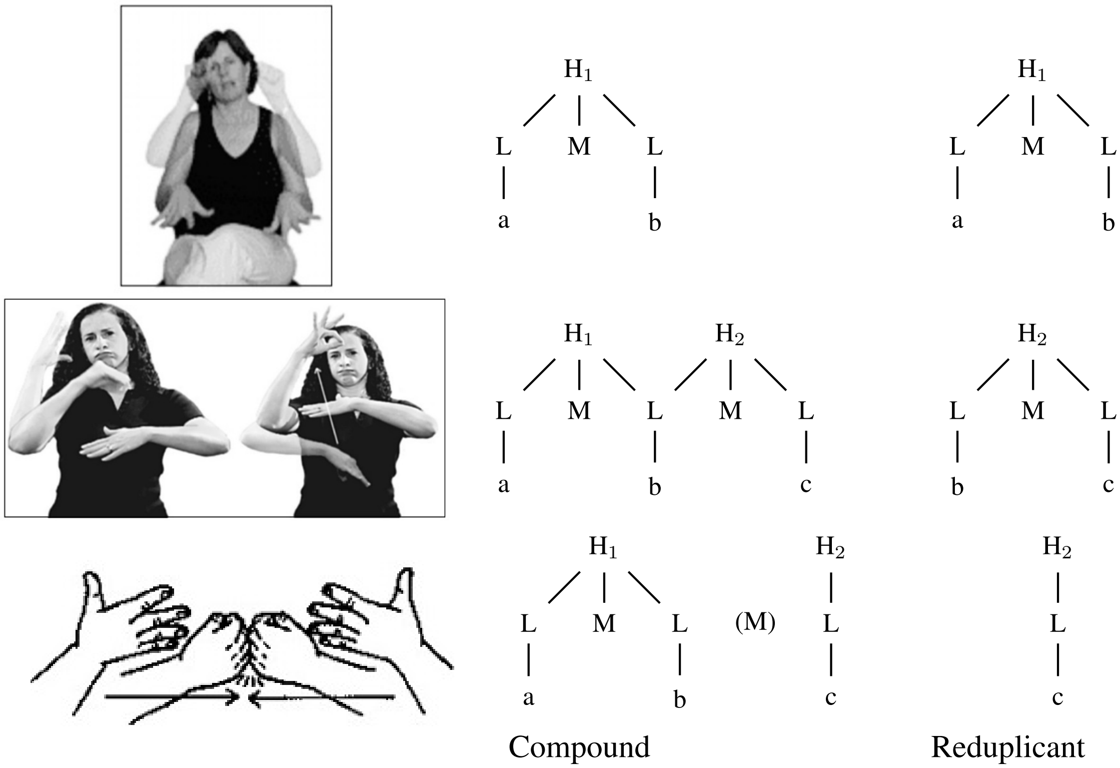


Figure 12: Final-syllable reduplication in ASL compounds FAINT (top), OVERSLEEP (middle), and ACCIDENT (bottom). Images Copyright Wendy Sandler and Diane-Lillo-Martin

(2) Tagalog (Blake, 1917)
súlat ‘write’ → su-súlat ‘will write’
$\emptyset \rightarrow C_1V_1 \setminus \#\_C_1V_1$

(3) Marshallese (Byrd, 1993)
ebbok ‘to make full’ → ebbok-bok ‘puffy’
lɣŋɔŋ ‘fear’ → lɣŋɔŋ-ŋɔŋ ‘very afraid’
$\emptyset \rightarrow C_1V_1C_2 \setminus C_1V_1C_2\_\#$

Chandlee (2017) shows that all local reduplication patterns are Input Strictly Local, since the partial information copied from the input is attached to the same edge it is copied from. She contrasts this with non-local reduplication, where the copied portion attaches to the opposite edge from where it is copied, and shows that this is not an ISL function.

As an example, consider the reduplicated disyllabic compound ‘OVERSLEEP’, which copies the second LML sequence, along with the second handshape and the associated place features. As above we can use the same model signature here, and the word model for ‘OVERSLEEP’ is shown visually in Figure 13.

This partial copying process takes advantage of the model-theoretic ability to specify multiple copies of the input structure, known as copy sets. To capture the reduplicated element, I define two copies of the input domain, shown visually in Fig. 13. Each copy keeps all information about the input structure: each labeling and association relations are the identity relation (omitted for space),

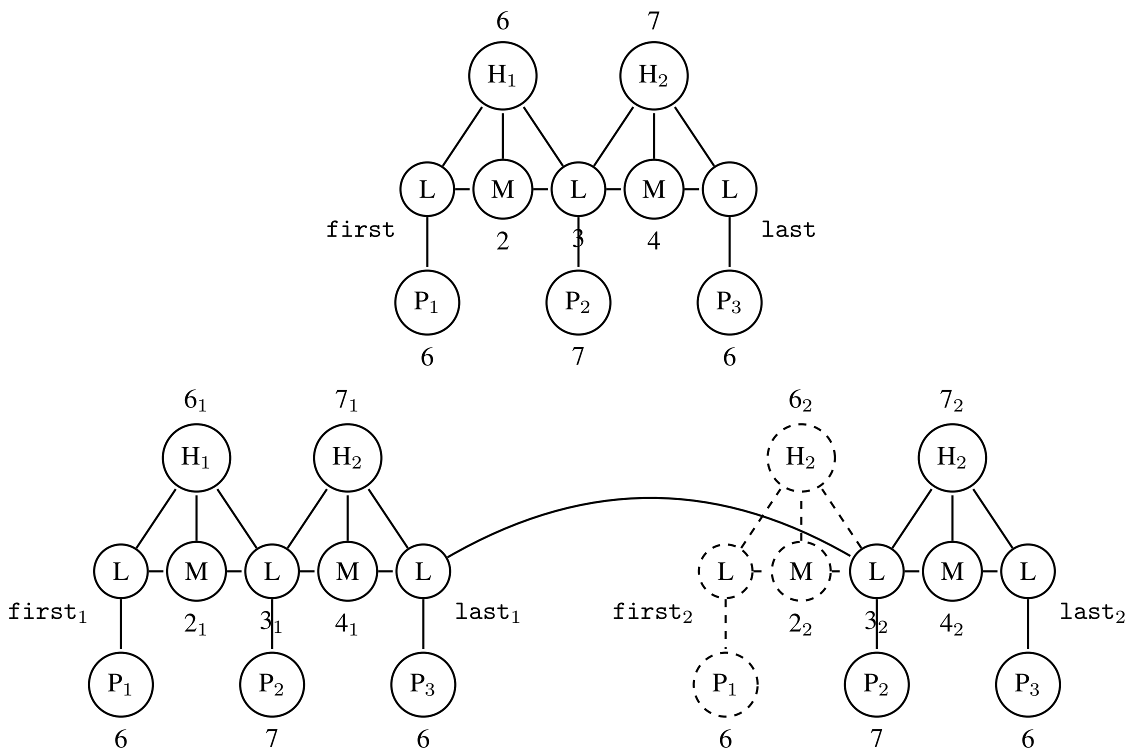


Figure 13: Visual of Input Model (top) and Output Copy Sets (bottom left and right) for Reduplicated ASL 'OVERSLEEP'. Unlicensed elements are dashed.

and all elements are licensed, as shown in (24), where the the subscript denotes the corresponding copy set. (25) says that the final, penultimate, and antipenultimate LM-tier segments are preserved, and that only the final handshape defined for the `last` segment is preserved. This captures the 'second syllable' reduplication copy.

(24) $$lic_1^O(x) \overset{\text{def}}{=} \texttt{TRUE}$$

(25) $$\begin{aligned} lic_2^O(x) \overset{\text{def}}{=} & [(L(x) \vee M(x)) \wedge x = [\texttt{last} \vee \texttt{p}(\texttt{last}) \vee \texttt{p}(\texttt{p}(\texttt{last}))]] \\ & \vee [\bigvee_i H_i^O(x) \wedge A(x, \texttt{last})] \\ & \vee [\bigvee_i P_i^O(x) \wedge x = [\texttt{loc}(\texttt{last}) \vee \texttt{loc}(\texttt{p}(\texttt{p}(\texttt{last})))]] \end{aligned}$$

To join the two copy sets and make a reduplicated word, we must specify the successor and predecessor functions for each copy set, and between copies. In (26-27), (x, i) refers to the ith copy of element x in the output structure, and in Figure 13 this is shown as x_i. (26) defines the successor function in a particular copy set i. The function says that for all non-final elements in the first and second copies, their successor is the same as in the input. For the final element in copy 1, its successor is the antipenultimate element of copy 2. Similarly, the `pred` function in (27) says that precedence in the output sets is the same, except for the antipenultimate segment of copy 2, whose predecessor is the final element of copy 1. A visual representation of the final input structure is provided in Figure 13. If the reduplicant were some other part of the input structure, then the licensing functions, and the successor and predecessor functions between copies 1 and 2 would change accordingly, without needing to change the copy-internal functions at all.

(26)
$$\mathtt{s}((x,i)) \stackrel{\text{def}}{=} \begin{cases} (\mathtt{s}(x),1), & x \neq \mathtt{last}, \quad i=1 \\ (\mathtt{p}(\mathtt{p}(\mathtt{last}),2), & x = \mathtt{last}, \quad i=1 \\ (\mathtt{s}(x),2), & i=2 \end{cases}$$

(27)
$$\mathtt{p}((x,i)) \stackrel{\text{def}}{=} \begin{cases} (\mathtt{p}(x),1), & i=1 \\ (\mathtt{last},1), & x = \mathtt{p}(\mathtt{p}(\mathtt{last}), i=2 \\ (\mathtt{p}(x),2), & x \neq \mathtt{p}(\mathtt{p}(\mathtt{last}), i=2 \end{cases}$$

To conclude, this section presented a model-theoretic characterization of several phonological processes in sign language. Each of the processes was sufficiently described by the quantifier-free logical transformations over autosegmental representations, meaning each of the mappings is A-ISL. These results extend those of Rawski (2017), showing that moving from strings to a richer representational system still preserves the strictly local nature of these processes.

Of course this is not an exhaustive description of sign language phonology. Sign languages regularly display assimilation (total and partial), deletion, epenthesis, reduction, coalescence, and many other processes (see Gu (2018) for an overview of these in Shanghai Sign Language). This section showed how the model-theoretic perspective can be used to investigate these and determine the precise nature of any sign language process. One only needs to be explicit about the representations and relations that characterize the structures and transformations between them. However, the bounded nature of the sign representation makes a strong prediction that they will be A-ISL. For instance, assimilation is never unbounded in or between signs, reduplication will only ever make reference to a bounded amount of material, and reduction involves bounded material. Strict Locality appears to be representationally salient for phonological computation across modalities. This provides some intriguing consequences for a longstanding debate in phonology on the relationship between sequentiality and simultaneity across modalities.

5 Locality, Sequentiality, and Simultaneity

This section examines the nature of sequentiality and simultaneity in light of the Strict Locality discussed above. Languages in both modalities have sequential structure, but there are striking differences in the nature of that structure, as the model-theoretic perspective shows. Spoken languages vary in syllable structure, word length, and stress patterns among syllables. Sign languages appear limited in all these aspects. They are overwhelmingly monosyllabic, have no clusters, and show extremely simple stress patterns, due to few polysyllabic words apart from fully reduplicated forms (Wilbur, 2011).

As the previous sections showed, the structural organization in signed or spoken language has a direct effect on the phonology. Strikingly few phonological rules that are not morphosyntactically triggered have been discovered in sign languages, mainly due to sign's lack of sequential structure. Significant sequential structure in sign mainly appears under morphosyntactic operations that concatenate morphemes and words (affixation, compounding, and cliticization). Predictably, when this occurs, a smorgasbord of phonology arises.

In general, sequential affixation is rare across sign languages (Sandler, 1996; Aronoff et al., 2005), and sign exhibits a strong tendency to express concatenative morphology through compounding (Meir, 2012). Aronoff et al. (2005) show that affixation usually results from grammaticalization of free words, via a series of diachronic changes concerning phonological and semantic factors. They cite the relative youth of sign languages as causing their lack of affixes compared with verbal languages. No known sign languages are over 300 years old, with some like Nicaraguan Sign Language, as young as 40 (Woll et al., 2001).

The lack of sequential structure in sign languages does not imply structural simplicity, however. Sign languages routinely employ nonconcatenative morphology (Emmorey, 2001; Meier, 2002), incorporating morphological material simultaneously in the phonology with restricted sequential form. Of course, simultaneous phonological structure exists in all languages, but differ across modalities in the amount. Very few sign features actually become sequenced, while in spoken language features are overwhelmingly sequenced, rarely simultaneous.

Comparing hand configuration and place autosegments to autosegments in spoken language shows further differences. Unlike spoken autosegments for tone or harmony patterns, which typically consist of one or two features, the hand configuration autosegment in sign languages is extremely complex, containing almost half of sign features organized in an intricate feature geometry (van der Hulst, 1995; Sandler, 1996). Also, only a subset of spoken languages utilize long-distance autosegments, while sign language is arguably universally characterized by autosegmental hand configuration and place of articulation.

The model-theoretic perspective highlights an intriguing possibility. The logical and computational nature of tonal patterns in spoken languages is among the most computationally powerful in spoken phonology. Jardine (2016a) shows that tonal patterns with string representations require the full power of the regular relations, or equivalently, full Monadic Second-Order Logic. In contrast, all known segmental processes occupy subclasses of these relations. Jardine conjectures that tonal patterns are by nature computationally more expressive. Jardine (2017b), and Chandlee and Jardine (2019) show that representationally switching to graphs that capture the autosegmental nature of tone pushes the complexity of some processes down, but some are still beyond A-ISL.

If sign languages were to take full advantage of the simultaneous abilities that the modality provides, and at the same time have access to the sequential range that spoken processes have, the complexity of the linguistic processes might dramatically increase, out of the ISL class, and perhaps even out of the regular class. A similar claim could be made for spoken languages if they had access to the same simultaneous aspects that sign has.

Alternatively, if the phonological system is computationally limited to, or at least strongly biased toward Strict Locality in its mappings, as I conjectured earlier, then a representational tradeoff arises naturally. Spoken languages, limited in their simultaneity but not their sequentiality, can satisfy the computational ISL requirements (bounded locality and a limited memory) by expressing the majority of their phonology sequentially, and limiting the simultaneous expressivity to certain circumstances. Conversely, sign languages may limit the amount of sequentiality in phonological operations, satisfying the computational requirements via the representation by taking advantage of the rich simultaneity afforded them by the nature of the modality and the articulation system. The resulting small number of sequential distinctions in signs may also be compensated for by a larger number of features, to maintain a similar number of lexical contrasts as spoken language.

While the capacity for linearity and non-linearity are common to both spoken and signed languages, the relative computational centrality of each differs in the phonological organization of each modality. This presents a more nuanced version of the “adapted system" view of sign language in humans specialized by evolution for use of spoken language. If aspects of the articulatory/perceptual system are somehow compromised, the computational pressures characteristic of phonological organization combined with the modular features of the sign modality enable representations where simultaneity takes power over sequentiality.

6 Phonology, Cognition, and Modality

David Poeppel (2012) identifies a central challenge of the cognitive neuroscience of language to “identify representations and operations that can be linked to the types of operation that simple electrical circuits can execute." He suggests “theoretically well-motivated units of representation or processing deriving from cognitive science research (here, say linguistics); then one attempts to decompose these into elementary constituent operations that are formally generic". This is exactly what Model Theory provides. Model-theoretic representations of linguistic structure, and the logical information needed to compute the constraints and processes characterizing human language, make principled claims about the nature of cognitive representations and operations that underlie them. This has direct implications for the interplay between phonology and modality.

Recall that phonology across speech and sign is sufficiently characterized by Regular languages and relations, which are describable with MSO logic and finite-state machinery. Any cognitive mechanism that can perform computations over a finite-state set must be capable of classifying input-output members into a finite set of abstract categories and be sensitive to the sequence of those categories (Rogers et al., 2013). This subsumes any processing mechanism in which the amount of information inferred or retained is limited by a fixed finite bound. Any cognitive mechanism that has such a fixed finite bound in processing sequences of events will be able to recognize only finite-state sets of structures or processes.

Across spoken and signed processes, “regularity" is a sufficient, but not necessary, condition on the cognitive capacity required for processing. The restriction to or bias toward processes describable by QF mappings or ISL functions is a stronger suffucient claim, as these are a proper subset of the Regular relations. Any cognitive mechanism that can distinguish input-output relations of this kind must be sensitive, at least, to a bounded number of blocks of events that occur in the presentation of the structure. The sufficiency of local evaluation inherent to QF logic in computing the functions characteristic of natural language phonology is important. It shows that the phonological module is amodally sensitive to bounded chunks of structural information at each moment of the process. The model-theoretic perspective makes this clear, by precisely defining the nature of the structures and the computation working over them. In more cognitive terms, this means looking at finite chunks of structure and a very restricted notion of memory.

This is not to say that the discovery of a process in either modality that is not describable by a QF mapping, or an ISL function, over graphs, invalidates the claims. In fact, several have been discovered, though with increasing dispreference as one approaches full regular power (see Heinz (2018)). What it does say is that one has choices. If one is committed to a cognitive view where computation is limited and works in this local way, then one can choose to impose more structure into the model signature, or adjust the structures which are already there. This is the approach taken by Heinz (2010), who makes the case for a general precedence relation rather than a successor relation, and by Jardine (2017a), who argues for incorporating more structure to handle autosegments.

Alternatively, if one is committed to the view of phonology with a particular structure, meaning a particular signature or part of a signature, then one can understand the relationship of that structure to the range of computational power needed to handle phonological processes using that structure . Most of the work in understanding the subregular complexity of phonological processes has assumed a string structure, and there is extensive work carving out the precise range of computations required for phonology (Graf, 2017). This too has consequences, since this carving may require a further modular view of the phonological system depending on the classes of computations one finds to be necessary and sufficient. This also modularizes the view of the learning system, as successful learning algorithms are often tied to the data structures which characterize these function classes (Heinz, 2010; Heinz and Idsardi, 2013).

The salience of certain representations and computational aspects across modalities suggests that certain parts of the phonological module are amodal, that is, independent of modality, and in some cases constrain representation in the specific modality the phonology is expressed in. So where does this leave modality effects? Model theory allows explicit comparison of the nature of phonological properties across modalities to see where these differences lie: representation, or computation? If a particular process differs across modalities, one can precisely characterize which aspect of representation or computation is responsible. This gives a promising avenue for future research on modality and amodality in the phonological system, and concrete testable hypotheses for experimental approaches.

As mentioned in the preceding section, the phonological module may accommodate the representational abilities of the particular articulatory/perceptual system to satisfy the requirements of computation. This view has some independent support. van der Hulst and van der Kooij (2018), cite Brentari (2002) and Emmorey (2001) that visual perception of signs (even with sequential properties) is more “instantaneous" than auditory speech perception, and adapt Goldsmith (1976)’s division of phonology in terms of the notions of “vertical and horizontal slicing of the signal". They state

> an incoming speech signal is first spliced into vertical slices, which gives rise to a linear sequence of segments. Horizontal slicing then partitions segments into co-temporal feature classes and features. In the perception of sign language, however, the horizontal slicing takes precedence, which gives rise to the simultaneous class nodes that we call handshape, movement, and place. Then, a subsequent vertical slicing of each of these can give rise to a linear organization

Of course, the phonetics and phonology of sign language differ in many ways, and this isn’t surprising. Lillo-Martin (1997) cites Blakemore (1974)’s result that exposure to vertical and horizontal lines in the environment affects development of feline visual perception, and asks “why shouldn’t exposure to the special acoustic properties of the modality affect perception, especially auditory perception?" Sandler and Lillo-Martin (2006) note that unlike spoken syllables in many languages, sign language syllables prohibit location clusters comparable to consonant clusters, or diphthong-like movement clusters, and there must be a movement between locations due to the physiology of the system. Additionally, sign syllables do not have onset-rhyme asymmetries, which affects syllable structure and stress assignment, they typically align intonation, conveyed by facial expression, with phonological/intonational phrases, not syllables inside those phrases, where spoken languages usually do (Nespor and Sandler, 1999).

However, the fact remains that locality and a bounded memory are representationally salient for and computationally exploited by the phonological module across modalities. These properties constrain the expression of phonological content by taking advantage of the particular properties of the modality. This represents a sufficient condition for amodality, and offers a promising route to exploring other necessary and sufficient conditions for it. The freedom and preciseness given by model-theoretic phonology and the computational tradeoffs that come with it give a promising answer to Poeppel (2012)'s call to "focus on the operations and algorithms that underpin language processing", since "the commitment to an algorithm or computation in this domain commits one to representations of one form or another with increasing specificity and also provides clear constraints for what the neural circuitry must accomplish."

7 Conclusion

This article presented a logical characterization of the nature of phonological properties in spoken and signed language. Model-theoretic analyses of phonological processes in sign were shown to require the same logical power as their spoken counterparts, namely, quantifier-free mappings, or ISL functions. It was predicted that almost all phonological processes in sign share this complexity due to the bounded nature of the sign, just as most phonological processes in spoken language fall into the ISL class. It was further conjectured that this computational constraint causes a tradeoff in the organization of phonological representations in each modality — more sequential structure in speech, and more simultaneous structure in sign. This has strong implications for the nature of the phonological module as an aspect of the cognitive capacity for language, highlighting the relevance of model-theoretic methods in addressing representational and cognitive questions, and providing a principled way to investigate the nature of language across speech and sign.

References

Aronoff, M., Meir, I., and Sandler, W. (2005). The paradox of sign language morphology. *Language*, 81(2):301.

Berent, I. (2013). *The Phonological Mind*. Cambridge University Press.

Blake, F. R. (1917). Reduplication in tagalog. *The American Journal of Philology*, 38(4):425–431.

Blakemore, C. (1974). Developmental factors in the formation of feature extracting neurons. *The neurosciences: Third study program*, pages 105–113.

Brentari, D. (1998). *A prosodic model of sign language phonology*. Mit Press.

Brentari, D. (2002). Modality differences in sign language phonology and morphophonemics. *Modality and structure in signed and spoken languages*, pages 35–64.

Büchi, J. R. (1960). Weak second-order arithmetic and finite automata. *Mathematical Logic Quarterly*, 6(1-6):66–92.

Byrd, D. (1993). Marshallese suffixal reduplication. In *Proceedings of the Eleventh West Coast Conference on Formal Linguistics*, pages 61–77.

Chandlee, J. (2014). *Strictly Local Phonological Processes*. PhD thesis, U. of Delaware.

Chandlee, J. (2017). Computational locality in morphological maps. *Morphology*, pages 1–43.

Chandlee, J., Eyraud, R., and Heinz, J. (2014). Learning strictly local subsequential functions. *Transactions of the Association for Computational Linguistics*, 2:491–503.

Chandlee, J. and Heinz, J. (2018). Strict locality and phonological maps. *Linguistic Inquiry*.

Chandlee, J., Heinz, J., and Jardine, A. (2018). Input strictly local opaque maps. *Phonology*, 35(2):171–205.

Chandlee, J. and Jardine, A. (2019). Autosegmental input strictly local functions. *Transactions of the Association for Computational Linguistics*.

Chandlee, J. and Lindell, S. (2016). Local languages. Paper presented at the 4th Workshop on Natural Language and Computer Science, in affiliation with LICS at Columbia University, NY.

Chomsky, N. (1956). Three models for the description of language. *IRE Transactions on information theory*, 2(3):113–124.

Courcelle, B. (1994). Monadic second-order definable graph transductions: a survey. *Theoretical Computer Science*, 126(1):53–75.

Davidson, J. O. (1977). *A Contrastive Study of the Grammatical Structures of Aymara and Cuzco Kechua*. PhD thesis.

Emmorey, K. (2001). *Language, cognition, and the brain: Insights from sign language research*. Psychology Press.

Enderton, H. B. (2001). *A Mathematical Introduction to Logic*. Academic Press, 2nd edition.

Engelfriet, J. and Hoogeboom, H. J. (2001). MSO definable string transductions and two-way finite-state transducers. *ACM Transactions on Computational Logic (TOCL)*, 2(2):216–254.

Fagin, R., Stockmeyer, L., and Vardi, M. (1995). On monadic NP vs monadic co-NP. *Information and Computation*, 120(1):78 – 92.

Filiot, E. (2015). Logic-automata connections for transformations. In *Indian Conference on Logic and Its Applications*, pages 30–57. Springer.

Filiot, E. and Reynier, P.-A. (2016). Transducers, logic and algebra for functions of finite words. *ACM SIGLOG News*, 3(3):4–19.

Finley, S. (2009). *Formal and cognitive restrictions on vowel harmony*. PhD thesis, The Johns Hopkins University.

Frishberg, N. (1975). Arbitrariness and iconicity: historical change in american sign language. *Language*, pages 696–719.

Goldsmith, J. A. (1976). *Autosegmental phonology*, volume 159. Indiana University Linguistics Club Bloomington.

Graf, T. (2010). Logics of phonological reasoning. Master's thesis, University of California, Los Angeles.

Graf, T. (2017). The power of locality domains in phonology. *Phonology*, 34(2):385–405.

Gu, S. (2018). *The Feature System of Hand Shapes and Phonological Processes in Shanghai Sign Language*. PhD thesis, Shanghai: East China University.

Heinz, J. (2010). Learning long-distance phonotactics. *Linguistic Inquiry*, 41(4):623–661.

Heinz, J. (2018). The computational nature of phonological generalizations. In Hyman, L. and Plank, F., editors, *Phonological Typology*, Phonetics and Phonology, chapter 5, pages 126–195. De Gruyter Mouton.

Heinz, J. and Idsardi, W. (2013). What complexity differences reveal about domains in language. *Topics in cognitive science*, 5(1):111–131.

Jardine, A. (2016a). Computationally, tone is different. *Phonology*, 33(2):247–283.

Jardine, A. (2016b). *Locality and non-linear representations in tonal phonology*. PhD thesis, University of Delaware.

Jardine, A. (2017a). The expressivity of autosegmental grammars. *Journal of Logic, Language and Information*, pages 1–46.

Jardine, A. (2017b). The local nature of tone-association patterns. *Phonology*, 34:385–405.

Johnson, C. D. (1972). *Formal aspects of phonological description*. The Hague: Mouton.

Kaplan, R. M. and Kay, M. (1994). Regular models of phonological rule systems. *Computational linguistics*, 20(3):331–378.

Klima, E. S. and Bellugi, U. (1979). *The Signs of Language*. Harvard University Press.

Lepic, R. (2015). *Motivation in morphology: Lexical patterns in ASL and English*. PhD thesis, UC San Diego.

Libkin, L. (2004). *Elements of Finite Model Theory*. Springer.

Liddell, S. K. (1984). Think and believe: sequentiality in american sign language. *Language*, pages 372–399.

Liddell, S. K. and Johnson, R. E. (1989). American sign language: The phonological base. *Sign language studies*, 64(1):195–277.

Lillo-Martin, D. (1997). The modular effects of sign language acquisition. *Relations of language and thought: The view from sign language and deaf children*, pages 62–109.

Meier, R. P. (2002). Why different, why the same? explaining effects and non-effects of modality upon linguistic structure in sign and speech. *Modality and structure in signed and spoken languages*, pages 1–25.

Meir, I. (2012). Word classes and word formation. *Sign language. An international handbook. Berlin: De Gruyter Mounton*, pages 77–112.

Mielke, J. (2007). P-base 1.92. *Available (July 2010) at http://aix1. uottawa. ca/˜ jmielke/pbase/index. html*.

Nespor, M. and Sandler, W. (1999). Prosody in israeli sign language. *Language and Speech*, 42(2-3):143–176.

Perlmutter, D. M. (1993). Sonority and syllable structure in american sign language. In *Current issues in ASL phonology*, pages 227–261. Elsevier.

Poeppel, D. (2012). The maps problem and the mapping problem: two challenges for a cognitive neuroscience of speech and language. *Cognitive neuropsychology*, 29(1-2):34–55.

Potts, C. and Pullum, G. K. (2002). Model theory and the content of OT constraints. *Phonology*, 19:361–393.

Pullum, G. K. (2007). The evolution of model-theoretic frameworks in linguistics. In Rogers, J. and Kepser, S., editors, *Model-Theoretic Syntax at 10*, pages 1–10, Dublin, Ireland.

Rawski, J. (2017). Phonological complexity across speech and sign. In *Proceedings of the 53rd Chicago Linguistics Society Annual Meeting*. (To appear).

Reiss, C. (2018). Substance free phonology. *The Routledge Handbook of Phonological Theory*, pages 425–452.

Rogers, J. (1998). *A descriptive approach to language-theoretic complexity*. CSLI Publications Stanford, CA.

Rogers, J., Heinz, J., Fero, M., Hurst, J., Lambert, D., and Wibel, S. (2013). Cognitive and subregular complexity. In Morrill, G. and Nederhof, M.-J., editors, *Formal Grammar*, volume 8036 of *Lecture Notes in Computer Science*, pages 90–108. Springer.

Rogers, J. and Pullum, G. K. (2011). Aural pattern recognition experiments and the subregular hierarchy. *Journal of Logic, Language and Information*, 20:329–342.

Sandler, W. (1986). The spreading hand autosegment of american sign language. *Sign Language Studies*, 50(1):1–28.

Sandler, W. (1989). *Phonological representation of the sign: Linearity and nonlinearity in American Sign Language*, volume 32. Walter de Gruyter.

Sandler, W. (1993a). Linearization of phonological tiers in asl. In *Current Issues in ASL Phonology*, pages 103–129. Elsevier.

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

Sandler, W. (1996). *Representing handshapes*, volume 1. Lawrence Erlbaum Associates, Inc., Mahwah, NJ.

Sandler, W. and Lillo-Martin, D. (2006). *Sign language and linguistic universals*. Cambridge University Press.

Schlenker, P. (2018). Visible meaning: Sign language and the foundations of semantics. *Theoretical Linguistics*, 44(3-4):123–208.

Scobbie, J. M., Coleman, J. S., and Bird, S. (1996). *Key aspects of Declarative Phonology*. European Studies Research Institute, Salford.

Shoenfield, J. R. (1967). *Mathematical logic*, volume 21. Reading: Addison-Wesley.

Strother-Garcia, K. (2018). Imdlawn tashlhiyt berber syllabification is quantifier-free. In *Proceedings of the Society for Computation in Linguistics*, volume 1.

Thomas, W. (1997). Languages, automata, and logic. In *Handbook of Formal Languages*, volume 3, chapter 7. Springer.

van der Hulst, H. (1993). Units in the analysis of signs. *Phonology*, 10(2):209–241.

van der Hulst, H. (1994). Dependency relations in the phonological representation of signs. *Sign language research*, pages 11–38.

van der Hulst, H. (1995). The composition of handshapes. *Trondheim Work. Papers*, 23:1–17.

van der Hulst, H. and van der Kooij, E. (2018). Phonological structure of signs – theoretical perspectives. In Josep Quer Villanueva, R. P. and Herrmann, A., editors, *The Routledge Handbook of Theoretical and Experimental Sign Language Research*. Routledge.

Wilbur, R. (2011). Sign syllables. *The Blackwell companion to phonology*, 1:1309–1334.

Williams, E. S. (1976). Underlying tone in margi and igbo. *Linguistic Inquiry*, pages 463–484.

Woll, B., Sutton-Spence, R., and Elton, F. (2001). Multilingualism: The global approach to sign languages. *The sociolinguistics of sign languages*, pages 8–32.