Abstract. Personal, possessive, and deictic pronouns in Israeli Sign Language cliticize onto hosts in ways that are determined by the syntactic structure as well as the prosodic structure of the utterance. These cliticizations show that the very different phonetic system of sign languages subserves a prosodic structure that is comparable to that of spoken languages, providing novel evidence for universals of prosodic structure. The two types of cliticization examined adhere in some ways to the optimal form of the prosodic word in sign languages. A constraint based analysis is suggested, to account for the fact that many complex words as well as host plus clitic constructions obey surface well-formedness constraints on canonical simple signs. The study suggests that the interaction of well-formedness constraints is universal to all human language, although the constraints themselves are not.

1. Introduction.

Sign language phonology is perhaps the most lively area of theoretical linguistic research on sign languages, and has been so, ever since William Stokoe (1960) first demonstrated that there is such a thing. There are two reasons for this intense interest in the phonological level of structure. One is that the very existence of a phonology -- a finite list of meaningless units that combine in constrained ways to form meaningful lexical contrasts -- shows that sign languages are characterized by duality

---

1 I am grateful to the participants of the conference on the Phonological Word in Berlin, at which this work was first presented, for their helpful comments and questions. Marina Nespor and Ellen Broselow offered useful comments on earlier versions of this paper, for which I thank them. I am especially indebted to Laura Downing for her astute and constructive comments and discussion of this paper. Mistakes are mine alone. Thanks also to the project research assistant, Irit Meir, and sign language consultants Meir Etedgi, Orna Levy, and Doron Levy for their important contributions to this research. This research is funded by the Israel Science Foundation, grant no. 820/95, and by the Binational Science Foundation, grant no. 9500310/2.
of patterning, a defining property of human language. The discovery of this property made it impossible to dismiss sign languages as a collection of iconic, holistic gestures that are inherently different from the words of spoken languages. The other reason that the phonological level is so intriguing is that it is so intimately bound to the physical production system. Since this physical, phonetic, system of sign languages is fundamentally different from that of spoken languages, one might expect the structure and organization of the phonological system it articulates to be fundamentally different as well. Yet, we have been finding significant similarities.2

This study is concerned with one aspect of the interaction between phonology and syntax. The results reported here are part of a larger project on prosody and intonation in Israeli Sign Language (ISL), which has demonstrated the existence of higher levels of prosodic structure than the one dealt with here, namely, the phonological phrase and the intonational phrase (Nespor and Sandler 1997, to appear). The present study examines two different ways in which pronouns merge phonologically with hosts, each in different prosodic environments within the phonological phrase. There are three reasons for claiming that the merging is postlexical (in the sense of Kiparsky 1982, Kaisse and Shaw 1985, Booij 1994, 1997). The most obvious is that two morphosyntactic words — units determined by the syntax — blend together in certain positions within the phonological phrase, (Nespor and Vogel 1986). Second, the processes are optional but have no lexical exceptions. Third, the resulting forms are non-structure preserving.3

The mergings demonstrate that certain constraints on the form of prosodic words are active postlexically, where function words cliticize to hosts. At the same time, some constraints on prosodic words are violated in each type of merging reported here. This raises the suggestion that constraints may be ranked differently at the lexical and postlexical levels


3 In their paper, American Sign Language and the Architecture of Phonological Theory, Padden and Perlmutter give evidence for a distinction between lexical and postlexical rules. This distinction is also argued for in Sandler (1993c).
3

(Booij 1997). I leave a formal Optimality Theory type analysis of these and other relevant sign language phenomena to ongoing work (Sandler, in preparation), and concentrate here on description and a less formal constraint oriented account. Constraints are posited — constraints that hold both at the lexical and at the postlexical level, though with different effects - and their interaction is described.

I begin with a brief description of the phonological structure of the sign, focusing on elements that are relevant to the topic under discussion here. Constraints on prosodic words that will be relevant for the clitics analysis are presented in Section 2, followed in Section 3 by a description of the two types of cliticization in Israeli Sign Language (ISL). I then go on to show that the two types of cliticization actually involve satisfaction of some of the same constraints that were shown in Section 2 to be active in the lexicon, while violating others. This state of affairs strongly suggests an analysis along the following lines: The same constraints are active lexically and post-lexically, but certain constraint rankings differ at the two levels. The basic insight that such an analysis reveals is that the clitic-host constructs are ‘trying’ to achieve the optimal form of the prosodic word. Section 4 compares constraints to rules for some of the phenomena dealt with. The next section discusses some unresolved issues, and the Conclusion highlights implications of this study for the issue of language universals.

2. The phonological structure of the sign: constraints on prosodic words

2.1. The canonical sign

The field of sign language phonology is relatively small and new, yet energetic, with several different (and often competing) models of sign language structure (presented, for example, in Liddell, 1984; Liddell and Johnson, 1989; Sandler, 1989, 1993a,b,c, 1996; Corina, 1989; Brentari 1990, to appear; Wilbur, 1993; Uyechi, 1996; van der Hulst, 1993, 1996). For overviews, I refer the reader to the references in footnote 2. I instantiate the structure using my own model (the Hand Tier model), focusing especially on those aspects of the model that bear on the present discussion, while citing different treatments where relevant. Much of the motivation for this model comes from American Sign Language (ASL). With respect to the phonology that is relevant to the present investigation, the same basic
phonological structure appears to characterize ISL (and sign languages in general), though more research is needed to confirm this.

Let us take a typical sign as an example. The picture in (1a,b) shows the Israeli Sign Language sign meaning roughly ‘to like’.

(1a) LIKE (beginning)  (1b) LIKE (end)

Signs are traditionally described in terms of three basic categories: hand configuration (consisting of handshape and palm orientation (Sandler 1987, 1989)), location, and movement. The handshape is further subdivided to selected fingers features, and specification of their position. In the ISL sign LIKE, the handshape is: all five fingers, in a curved position. Locations are comprised of a place of articulation -- the head, the trunk, or the nondominant hand -- as well as finer settings with respect to that place, such as ‘high’, ‘low’, etc. In LIKE, the signing hand articulates two settings which are in relation to the head place of articulation. The first setting of the hand at this place of articulation is in contact with it, and the second is at a medial distance from it. The movement is a default straight path from the first setting to the second. (2) is a schematic representation of a canonical sign, and (3) is a partial phonological representation of the sign LIKE. The sequential organization of locations and movements (Liddell 1984) is reflected in the model. In (3) I follow the convention established

4 The representation collapses certain categories for the sake of simplicity. In particular, the handshape category is made up of two subcategories not shown here: the selected fingers, and their position (Sandler 1989, 1995a, 1996a).

5 Liddell’s (1984) model and subsequent work with Johnson (e.g., 1989) proposes that the sequential units are holds (lack of movement) and movements. I assume that holds are not underlying (Sandler 1986, 1989) and that the relevant category is location instead.

6 Stokoe (1960) holds that the organization of the phonological categories of signs is simultaneous, unlike the sequential nature of spoken language segments. Subsequent work (e.g., Supalla and Newport 1978, Newkirk 1981, Padden 1983, Liddell 1984, Sandler 1989) provides evidence for sequential structure. The precise representation of this sequential structure, and whether or not it is underlyingly associated to particular phonological
in the Hand Tier model of representing the hand configuration hierarchy ‘upside down’.

(2) **schematic representation of a canonical sign**

![Diagram of a canonical sign]

(3) **example:** ISL LIKE - ‘be a fan of’

![Diagram of 'be a fan of']

This representation reflects certain constraints. While earlier analyses have proposed that these constraints hold either over the sign (= morphosyntactic word) or the morpheme, I propose that the relevant domain is the prosodic word. The rationale for this proposal will become evident in the following paragraphs.

The one-to-many association of hand configuration to the L and M segments illustrates the facts that certain aspects of hand configuration categories such as location and movement, is a subject of debate (cf.
remain constant throughout this domain, and that the category of hand configuration behaves autosegmentally (Sandler 1986, 1989). In particular, the Selected Finger Constraint requires that there be only one specification for choice of fingers per morpheme (Mandel 1981). I follow Brentari (in press) in suggesting that the appropriate domain for this constraint is the prosodic word. The constraint is formulated in (4).

(4) Selected Finger Constraint (SF): Only one specification for selected fingers is allowed in a prosodic word.

Brentari shows that this constraint holds over words containing one or more than one syllable and one or more than one morpheme, so that earlier formulations assuming those smaller domains are incorrect. This, together with the assumption that the morphosyntactic word is generally isomorphic with the prosodic word in ASL and ISL (to be discussed below), renders the constraint in (4). The hierarchical relationship of handshape and orientation in (3) is motivated by assimilation facts of reduced compounds in ASL (Sandler 1987, 1989; see footnote 25).

That there is also only one major body area (= place of articulation) in the prosodic word is also reflected in the representation (cf. Battison 1978 for a similar constraint, but on the ‘sign’). I refer to the latter requirement as the Place constraint.

(5) Place Constraint (P): Only one place of articulation may be specified in a prosodic word.

The SF constraint and the P constraint are always obeyed for native monomorphic words, and their effects are often felt on morphologically complex and borrowed forms as well, which is one reason for assuming the prosodic word domain. Another reason for assuming the prosodic word rather than the syllable as the domain for both the SF and P constraints is that they are satisfied in those few native ASL disyllabic monomorphic words that exist -- such as DESTROY. They do not hold between words, however. There are no known constraints on sequences of selected fingers or places of articulation in successive words.

The structure of the sign LIKE is canonical. Signs (i.e., words ) typically have only one timing unit characterized by movement, although different types of movement can be simultaneously superimposed on this timing unit. For example, some signs are characterized by a changing

handshape, in which the position of the fingers changes, e.g., from closed to open, or by a change in palm orientation. This ‘internal’ or ‘local’ movement is simultaneous with the path movement from one setting to another. It is generally assumed by those linguists who have posited the existence of a syllable-like unit in sign languages that the movement corresponds to the syllable nucleus (Coulter, 1982; Wilbur, 1982; Sandler, 1989, 1993a; Brentari, 1990; Perlmutter, 1992). While the internal movement resulting from a change in finger position or palm orientation may coincide with a path movement from one location to another, the simultaneous movements still constitute one syllable. Two movements in succession are counted as two syllables. This means that most monomorphemic words, and, as we shall see, many multimorphemic words, are monosyllabic. One may think of CVC as a heuristic comparison with the monosyllabic LML forms represented here. To account for this well known tendency for signs to surface as monosyllabic, I introduce the constraint in (6).

(6) MONOSYL. One syllable per prosodic word.
Prosodic words are monosyllabic.\(^7\)

While sign languages are typically complex morphologically, morphemes are generally not concatenated. Rather, they are usually integrated nonconcatenatively (Sandler 1989, 1993a,b), avoiding sequences of movements, and therefore not disturbing the canonical monosyllabic form. For example, the phonological realization of verb agreement morphemes in ASL, ISL, and other sign languages involves nonconcatenative association to the first and last locations of the sign, rendering the canonical LML form, as shown in (2).

\(^7\) Certain types of monomorphemic signs -- sign types that have largely predictable second syllables, such as lexically reduplicated signs -- may violate the monosyllable constraint, and it expected that the correct analysis of the redundancy will ultimately explain this.
‘I look at you’ ‘You look at me’

Figure 7. agreement in Israeli Sign Language

Similarly, temporal aspect morphology in both languages involves altering the quality or rhythmic structure of the movement, again without adding any concatenative morphology, and again resulting in monosyllabic bases (which may then reduplicate) (Sandler, 1990).

It seems that the sign language data reflect a kind of global conspiracy, in which morphological structure is competing with prosodic constraints and often losing (Sandler, 1993a, 1994, 1995b). That is, even when signs become morphologically complex, the output often conforms in certain ways to the simple prosodic form shown in (2) above.

2.2. Output constraints on complex words

While most morphological processes in ASL and ISL are nonconcatenative, there are some in which morphemes are added linearly. It has been shown for ASL that in some cases of linear affixation or compounding, the base signs are truncated, resulting in a monosyllabic (though bimorphemic) form on the surface.8 Examples are ‘negative incorporation’ (Woodward 1975), a nonproductive process that suffixes a negation marker, and lexicalized reduced compounds (Sandler 1993a,b).

Figure (8) shows the base sign WANT, and the form DON’T-WANT, with negative incorporation.9 Note that there is only one movement in each form.

---

8 Some limited true concatenative affixation has been reported in sign languages (e.g., Liddell, 1996, Sandler, 1996c, Aronoff, Meir and Sandler, in preparation). The negative marker discussed here should not be confused with the concatenative negative suffix described in the latter two references.

9 The term ‘incorporation’, from Woodward (1974), is misleading, since it obscures the fact that the negative affix always occurs at the end of the sign. It does, however, reflect the intuition that the negation marker is ‘incorporated’ into the canonical sign structure - LML in the HT framework.
In (9), the underlying and surface forms of a verb with negative incorporation are represented schematically. The output is formally identical to the canonical form of a prosodic word, shown in (2).

(9) base + negative marker-> attested form

```
L M L + L  -->  L M L
|   |   |   |   |
| a  b  palm down  b  palm down |
|   |   |   |   |
| hand open  hand open |
|   |   |   |   |
| low location  low location |
```

This process is suffixal: the negative marker (palm down, hand open, low location) is added after the last location of the base sign. If the negative marker were to be simply added linearly to the base, the result would be a disyllabic word. The last location of the base and the location of the marker are not the same (the marker is lower in the signing space), which would necessarily result in an epenthetic movement between them (Sandler, 1990). Rather than parsing the input straightforwardly, the attested form involves truncation of the first location, and the straight default movement to the second location. The location of the negative marker is added, resulting in epenthesis of a single movement, and yielding a monosyllabic output. This type of behavior is expected in a constraint based theory, such as Optimality Theory (Prince and Smolensky, 1993; McCarthy and Prince, 1993), according to which the output differs from the input because it obeys general constraints on surface form. In this case, the Monosyllabicity constraint is ranked above a constraint requiring faithfulness to the input, such as MAX (McCarthy and Prince, 1995).
As I have said, compounds also may reduce to the optimal prosodic word form. Figure (10) illustrates the lexicalized ASL compound FAINT, made up of the two signs MIND and DROP. The two members of the compound, each monosyllabic in citation form, are truncated, resulting in a compound which is monosyllabic rather than disyllabic. Such reduction in lexicalized compounds occurs in ISL as well. Fewer examples have been found in ISL so far, which may be related to the fact that the language is a lot younger.

Figure (10). The ASL compound FAINT, made up of MIND and DROP

We may schematize this reduction, in which two syllables become one, as in (11).

(11)

\[
\begin{array}{c}
\text{HC1} \\
L1 \ M \ L2
\end{array}
+ \begin{array}{c}
\text{HC2} \\
L3 \ M \ L4
\end{array}
\rightarrow \begin{array}{c}
\text{HC2} \\
L1 \ M \ L4
\end{array}
\]

Notice again that the output is formally nearly identical to the canonical sign shown in (2). The lexicalized compound obeys two constraints that appear to hold over the prosodic word: Monosyllabicity and the selected finger constraint (SFC)\(^{10}\). Each individual member of the compound has a

\(^{10}\) Mandel's constraint had the 'sign' as its domain. Sandler's (1989) constraint posited the morpheme as its domain, because both compounds and the polymorphemic verbs of motion and location (VMLs) violate the constraint. However, if we allow that constraints are violable, then stating
different selected finger specification -- the index finger for MIND, and all five fingers for DROP. The lexicalized form has only the latter, and, in this way as well, conforms to the optimal form of the prosodic word. The effect of the Monosyllabicity constraint, stated in (6), is also observed in the lexicalized form, which has one syllable rather than the two of the input. However, the Place constraint is violated, as the attested form is articulated at the head and in front of the trunk.

The behavior of compounds is very complex, and is not the place to look for a clearcut distinction between the morphosyntactic word and the prosodic word. Nonlexicalized compounds may surface as two prosodic words, in that each member may independently obey constraints on prosodic words. In lexicalized compounds, the two members together may obey some constraints on prosodic words, as we have seen in FAINT, which obeys the SF and MONOSYL but not Place. This varied behavior makes it difficult to generalize over compounds in these sign languages with respect to their prosodic status. I am taking the tentative position here that the morphosyntactic word is usually isomorphic with the prosodic word in ASL and ISL. Compounds (and many other complex words) obey output constraints on prosodic words – some compounds and some constraints, but not all of them. On the assumption that constraints are violable, a constraint based theory offers promise for ultimately explaining this variable behavior. Our focus here is host plus clitic constructions, however, and I will argue in Section 3 that these forms also obey some of the constraints on prosodic words, but in a way of their own.

Let us return to the example FAINT, shown in (10), and consider the ranking of the constraints it obeys. In FAINT (and many other compounds), both Monosyllabicity and the SFC are obeyed, so that one can’t tell whether one of these two constraints is higher ranked than the other. However, other reduced compounds surface with only one syllable, but two selected finger specifications, and none that I know of surface with one SF and two syllables, suggesting the ranking MONOSYL > SF.

More evidence for this ranking (in ASL at least) comes from some lexicalized borrowings from fingerspelling. In American Sign Language, the domain for this constraint as the prosodic word is more explanatory. In any case, under the reformulation of the SFC as holding over the domain of the prosodic word, unreduced compounds may be expected to be larger than prosodic words and thus violate the constraint. As for VMLs, their morpho-
English words are sometimes represented by fingerspelling -- spelling the English word by a sequence of handshapes representing letters of the English alphabet. Many fingerspelled words have become lexicalized, and in the process, they tend to regularize towards the form of native signs (Battison, 1978). An example is the borrowing of 'job'. In the normal fingerspelling of this word, the first letter, 'J', is signed with the pinky extended, and rotating the orientation of the palm from facing outward to facing sidewards. In ordinary signs, this kind of rotation is 'sonorous' enough to constitute a syllable nucleus (Corina, 1990a; Brentari, 1990, 1993; Perlmutter 1992; Sandler, 1993a). The letter 'O' is formed by bringing all the fingers and thumb tips together, palm facing outwards, and 'B' involves extending all four fingers and folding the thumb across the palm, which is also facing outward.

(12a) J O B

In the borrowed form, the 'O' is omitted altogether, and the 'B' is articulated with the palm sidewards, in its position at the end of 'J'.

(12b) Borrowed sign, JOB

syntactic and prosodic structure are idiosyncratic in other ways as well, indicating that they require a special treatment in any case.
If the hand were to rotate outward again to produce a normal ‘B’, this would have the effect of creating an additional syllable, whose nucleus is the internal movement of hand rotation, following the initial rotation produced for ‘J’. But the hand does not rotate outward again, so that the lexicalized borrowed form has two selected finger specifications (‘J’ and ‘B’), but only one syllable. This, and many similar examples presented by Battison, support the suggestion that the ranking MONOSYL > SF is correct. We will see in the next section that cliticization facts suggest a reranking of these constraints postlexically.

2.3. Two-handed signs and the Symmetry Constraint

The last structural element of interest in the context of this investigation is the nondominant hand (normally the left hand in right-handed people and the right hand in left-handed people). The signs of sign languages may be either one-handed or two-handed. In a language that uses two anatomically identical, motorically independent elements, one might expect each to function as an independent articulator. If this were the case, then it would constitute a significant difference from spoken language, which involves one tongue tip, one tongue body, one velum, etc. However, this is not the case. Rather, in signs -- whether they are monomorphemic or morphologically complex -- the nondominant hand does not behave like an independent articulator.\[1\]

Two-handed signs are of two basic types; in each of these the nondominant hand -- let us call it h2 -- assumes a different role (Sandler, 1989, 1993c). Either it acts like a shadow articulator, assuming the same shape and articulating the same locations and movements as the dominant hand, or it behaves like an immobile place of articulation like the head or the trunk, and the dominant hand articulates at or near it.\[12\] In either case, the nondominant hand is not an independent articulator, and its existence does not motivate a proliferation of phonological categories.

---

1 There is a subsystem within the grammar of sign languages in which the two hands do have a significant degree of autonomy: the system of complex verbs of motion and location (Supalla 1982), in which each handshape may represent an independent classifier and combine with motion and location roots. Utterances within the VML system violate many other phonological constraints as well, and, for our purposes here, I consider that system separate and beyond the scope of this study.

12 This description is oversimplified for the sake of clarity. See the references in note 18 for more detailed descriptions.
Since the model adopted here provides a representation of the hand articulator as well as of places of articulation, there is no need to make any substantial changes or additions for the nondominant hand. This is desirable, as it reflects the fact that h2 functions either as part of the HC hand articulator -- in signs we will call double-handed -- or as one of the possible places of articulation -- in signs to be called hand-placed. Figure (13a,b) shows the ISL double-handed sign, SHOUT, and (13a,b) the ISL hand-placed sign, ALREADY (perfect marker). Figure (14a,b) shows how each sign type is represented. In (14), irrelevant details are omitted and handshapes are abbreviated with composite symbols. For simplicity, the abbreviation SF is used for hand configuration instead of the symbol HC shown in (2).

In the double-handed representation, each hand has its own node dominated (here) by SF; each is associated to the same handshape features
and, through the SF node, to the same location and movement features. The representation reflects the Symmetry Condition (Battison, 1978), here renamed the Symmetry Constraint (SYM). Rather than having the sign as its domain as Battison proposed, I assume once again that the correct domain is the prosodic word:

“If both hands of a sign (prosodic word/WS) move independently during its articulation, then both hands must be specified for the same location, the same handshape, the same movement ... , and the specifications for orientation must be either symmetrical or identical.” (p. 33)

(15) Symmetry Constraint (SYM): Align h2 as h1 within a prosodic word. (15) refers to double-handed signs, and appears to be undominated in prosodic words in ASL and ISL. In the hand-placed representation (14b), [h2] is the place feature, on a par with the [trunk] place feature of the double-handed representation, or the [head] place feature in figure (3) for the sign LIKE. The nondominant hand obeys different constraints in Hand-placed signs, which I will not discuss here.

In the reduced compound FAINT shown in (10), then, three constraints on the prosodic word are obeyed: Monosyllabicity, Selected Finger, and Symmetry. By deleting one nucleus, monosyllabicity is achieved. By spreading the hand configuration of the second member regressively and delinking the hand configuration of the first, SF is obeyed, and SYM as well, since the surviving hand configuration is a double-handed one, and the two hands behave symmetrically.

There is no evidence in these examples for a dominance relationship between SYM and MONOSYL. However, other evidence suggests that SYM dominates MONOSYL.13 Specifically, there is a candidate output which would require MONOSYL to dominate SYM, and this candidate does not occur. The candidate is a monosyllabic reduced form that is achieved by having the two hands articulate the two places of articulation simultaneously, over the same syllable duration. Consider for example a compound in which the first member is double-handed and the second one handed. A hypothetical form of the compound might be one in which each hand begins at the place specified for the first member of the compound, but the hands end up in two different places: h1 at the final location for the second member of the compound and h2 at the final location of the first member. While
compounds of this sort, in which the first member is double-handed sign and the second is one handed, have been observed in both ASL and ISL, such reduction has not been reported, presumably because such a form would violate the symmetry constraint. Let us hypothesize, then, that SYM dominates MONOSYL in the formation of compounds whose first members are double-handed and second members are one-handed. As we shall see below, such forms are not only phonetically possible, but actually occur postlexically, in clitics.

As far as the lexicon of sign languages is concerned, then, the fact that there are two anatomically identical elements does not create a phonological structure in which there are two identical but independent articulators, a structure that would be anomalous compared to that of spoken language. The reason that no such structure is called for is that in sign language lexicons, only one of these elements is an independent articulator: the dominant hand. In fact, where the nondominant hand behaves symmetrically, in double-handed signs, it freely deletes (Padden and Perlmutter 1987). However, the nondominant hand plays an interesting role at the prosodic level, participating in the delineation of the prosodic constituent, phonological phrase (see Nespor and Sandler 1997, to appear). Another way in which the behavior of the nondominant hand participates in prosodic structure is described in the next section.

---

13 A perspicuous transcription system is being developed that is intended to facilitate the construction of tableaux (Sandler, in preparation).
14 This analysis selects certain types of reduction in order to examine possibilities for constraint based analyses in sign language morphophonology, and is in no way intended to explain all of the compound reduction phenomena. For discussions of compound reduction see Sandler (1986, 1987, 1989, in preparation), Liddell and Johnson (1986), and Brentari (to appear).
15 Lexical contrasts that are minimally distinguished by the presence or absence of two hands are extremely rare. The pair, LIKE (one-handed) and INTERESTED (double-handed) are so contrasted in ASL, and TAKE and ADOPT in ISL. By far the usual case is that the nondominant hand in double-handed signs can optionally delete with no change in meaning.
16 The theory of the nondominant hand presented here and developed especially in Sandler (1993c) is not uncontroversial. Other theories hold that h2 should indeed be represented as a separate though dependent structure, and that there should be only one representation for h2 regardless of phonological role. A discussion of the issue here would take us too far from the main point of this article, but for a different view see Brentari and Goldsmith (1993), van der Hulst (1996), Brentari (to appear), and, for a discussion of the two views in one paper, van der Hulst and Sandler (1994).
3. Cliticized forms

The present investigation is part of a larger study of prosodic structure conducted together with Marina Nespor. The corpus for the study consists of thirty sentences translated from Hebrew into Israeli Sign Language by three native signers. The sentences were recorded on videotape and coded by a native signer and linguist research assistant. They were designed to determine the direction of branching in the syntax of this language, as well as patterns of prominence. We discovered the following correlates of rhythmic prominence in ISL, which appears to be a right-branching language: (1) Prominence is marked on the rightmost word of a phonological phrase, mainly by holding the hand still at the end of a sign and/or reduplicating the sign.17 (2) Prominence is marked on the rightmost word of an intonational phrase with size and rate of signing (large and slow is more prominent). Those results, which include novel evidence for the phonological phrase as a prosodic unit, are reported in detail in Nespor and Sandler (1997, to appear).

In the course of the investigation, a level of structure lower than the phonological phrase was also discovered, and it is this level that I report on here. In particular, two morphosyntactic words may optionally undergo phonological processes that have the effect of bonding them prosodically. The structure of the higher prosodic constituents is relevant for the present study because each of the two cliticization processes described here occurs in a different rhythmic position within the phonological phrase. Because this merging generally involves a prosodically weak function word and a syntactically related content word, and because the merging has the effect of further weakening the function word, I refer to both processes as cliticization. Neither the syntax nor the meaning can alone determine which of the two processes applies. Rather, it is the rhythmic position that seems to be the determining factor, implying that the process is prosodic. The forms conform to the definition of ‘simple clitics’ (Zwicky 1977, Anderson 1995 [1992]): “an element of some basic word class, which appears in a position relative to the rest of the structure in which the normal rules of the syntax would (or at least could) put it.” (Anderson, p. 200).

---

17 We show that this prominence pattern conforms to those of spoken languages, falling as it does on the right in a right-branching language (Nespor and Sandler 1997, to appear).
Although the phonetics of sign languages are completely different from spoken language phonetics, I will show that the system organizes these physical elements in a familiar way at the syntax-phonology interface. It will also be seen that borrowed elements from spoken Hebrew offer a window to the structure of prosodic words in ISL. Each type of cliticization is marked by a different phonological process, and each occurs in a different prosodic environment. The two phonological processes are: coalescence, which occurs in the prominent position of the phonological phrase; and handshape assimilation, which occurs in a weak position. In the course of the discussion, it will be shown that interpreting the phonological processes in terms of violable constraints renders an explanatory account of these phenomena.

It is likely that competing constraints on prosodic structure at the level of the phonological phrase and above are responsible for determining which of the two types of cliticization occurs. However, I will not attempt to provide that higher contextual account here. The goal here is to try to show how the constraints already motivated for prosodic words interact to produce each of the cliticized forms. The bigger picture, as well as a more detailed treatment of constraints active within the lexicon, are the subjects of ongoing research (Sandler, in preparation). The present analysis abstracts away from certain details which may ultimately require a finer inventory of constraints and concomitant interactions. But I believe that the approach and the overall results reported here are in the right direction.\(^\text{18}\)

3.1. Coalescence

In this process, the pronoun encliticizes to a double-handed host sign. In most cases, the host is a concrete noun and the clitic a deictic consisting of a pointing gesture and meaning, ‘there’. This kind of construction, ‘noun there’ is common in Israeli Sign Language which, like other sign languages, exploits reference points established in space for verb agreement and other grammatical relations (e.g., Lillo-Martin and Klima 1990, Meir 1998, to appear).

\(^{18}\) Shepard-Kegl’s (1985) analysis of the ASL lexicon evokes the notion of clitics. However, the concept of clitics assumed in that work is quite different from the one argued for here, so that a comparison must be considered beyond the scope of this article.
The form of the enclitization is as follows. Midway through the signing of the double-handed host, the dominant hand articulates the clitic. Simultaneously, the nondominant hand continues the signing of the noun, and the two hands complete the two signs at the same time. The result is precisely that candidate that does not occur under compounding, as discussed in section 2.2. The cliticized result is the reduction of two monosyllabic words to a single monosyllable. This cliticization normally occurs where the noun and deictic are at the right edge of a phonological phrase, a rhythmically prominent position. An example occurred in the sentence shown in (15). P stands for a phonological phrase boundary, and I stands for an intonational phrase boundary. The underlined sequence is the cliticized form.

(15)  \underline{shop there} ] P  \underline{around-corner} ] P  I  bankrupt ]I......

‘The shop around the corner went bankrupt.’

The plain form of the sign SHOP is shown in figure (16a,b). (17a,b) shows the beginning and end of the cliticized form SHOP-THERE, in which the nondominant hand (h2) is articulating the end of SHOP, and the dominant hand is articulating the end of THERE (which is normally a one-handed sign).
By coinciding with their hosts, these pronouns lose their syllabicacy, a phenomenon noted for example in English aux contraction (Selkirk 1984). As is the case for function words in spoken languages, deictic pronouns in other prosodic positions are normally rhythmically weak. It is possible that their appearance in the strong phonological phrase-final position is what triggers the coalescence, on the assumption that stress assignment can only ‘see’ lexical categories.\(^{19}\) Presumably, the coalescence of clitics in this position allows the syllable to which they are adjoined to receive the prominence required phrase finally.

The syntax of ISL (and of other sign languages as well) allows a good deal of movement, and there were a few examples in the corpus in which the clitic was not a deictic, but rather another type of pronoun, either a personal subject pronoun or a possessive pronoun. In all cases of coalescence, the pronoun follows the host in a prosodically strong position, even where such a position is syntactically unusual, as in the case of a subject pronoun which exceptionally follows a verb. This variation lends support to the claim that the cliticization process is essentially prosodic and not syntactic. The effect of coalescence is the creation of a monosyllabic form with much more nonredundant phonological information than usual -- possibly a heavy syllable -- in this rhythmically prominent position. The result has the optimal monosyllabic form of the prosodic word, with enough semantic and phonological weight to carry the prominence of the prosodic position it is in.\(^{20}\)

The result of coalescence is non-structure preserving, since, as we have seen, the symmetry constraint prevents the articulation of two different movements and locations with each hand simultaneously. In other words, the symmetry constraint is undominated lexically. Non-structure preserving phonological processes are typical of postlexical rules (Kiparsky 1982, Kaisse and Shaw 1985). As cliticization is necessarily post-lexical, this is expected. Yet, as will become clear, the coalesced form is the result of a constraint – Monosyllabic - whose effects are seen in the lexicon, and that is precisely the point.

\(^{19}\) Wilbur (to appear) observes that in American Sign Language, pronouns are not stressed phrase finally, while signs belonging to a lexical category receive prominence in that position.

\(^{20}\) The real phonological arguments that violation of symmetry results in heavy syllables are yet to be made; at least we may say that such violations – which result in two hands doing different things at the same time – create signs that are more complex.
Assuming that the many-to-one associations of the lexical signs shown in the figures above must conflate at some stage (Sandler 1993b for American Sign Language, following insights in McCarthy 1986), we may represent the phonetic form of the coalesced words schematically as shown here in (18).

(18)

\[
\begin{align*}
  h2x & \quad h2x & \quad h2x \\
  | & \quad | & \quad | \\
  a & \quad b & \quad c \\
  | & \quad | & \quad | \\
  L & \quad M & \quad L \\
  | & \quad | & \quad | \\
  a & \quad b & \quad d \\
  | & \quad | & \quad | \\
  h1x & \quad h1x & \quad h1y
\end{align*}
\]

In (18), a, b, c and d stand for conflated bundles of location features, the two hands are represented as h1 (dominant) and h2 (nondominant), and x and y stand for hand configuration features. The two hands are released from the lexical double-handed relation shown above in figure (14a), now enabling them to be characterized by different handshape features, and to articulate different locations and movements in this postlexical process. Over the first two LM timing slots (approximately), both hands articulate the same location, movement, and hand configuration features. On the last timing slot, each hand articulates a different bundle of location and hand configuration features. This is a clear violation of the Symmetry constraint.

Only double-handed signs, in which both hands articulate the syllable nucleus (movement), may be the host in this type of cliticization. The important thing to notice in the representation is that the coalesced sign is monosyllabic, consisting of three timing units, of which the second is the movement nucleus. The cliticized form, then, is obeying a constraint on optimal prosodic word structure: Monosyllabicity. Such forms occur postlexically but not lexically, as discussed in section 2.2. The explanation suggested here is that the constraints SYM and MONOSYL have the opposite
ranking lexically and postlexically. Lexically, the ranking appears to be
SYM > MONOSYL\(^{21}\); postlexically, it is MONOSYL > SYM.

The phonetics of this process would be impossible in spoken
language: the corresponding situation would require two tongues to
articulate two independent articulations, plus the ability to perceive the
two different places of articulation simultaneously. However, spoken
language **prosodic** equivalents have been widely attested, in which a
function word loses its syllabic status, and forms a single prosodic unit with
its host.

Another sign-language particular phenomenon offers a unique type
of evidence that the cliticized form is a single prosodic word in the mind of
the signer. The phenomenon is mouthing of Hebrew words. In ISL, signers
sometimes mouth the Hebrew translation of selected words in a sentence.
Certainly not all words are mouthed -- this would be impossible if only for
the reason that the syntactic and morphological structures of the two
languages are vastly different, so that mouthing of all words would be like
producing two different languages at the same time. Rather, only some
words are mouthed, according to a system that is not yet well understood, in
which the choice of words is apparently based on semantic, syntactic, and
possibly prosodic grounds. Normally, the timing of the mouthing coincides
with the signing of the corresponding sign language word: *Xanut* (‘shop’)
is mouthed over the same time span during which the sign is produced,
as can be seen (by lipreaders) in illustration (16a,b). While the mouthing is
less clearly discernable in our picture of the cliticized form (17a,b) than it is
in the actual spontaneous data, the cliticized forms involve mouthing of
only the word for ‘shop’ (and not the Hebrew word corresponding to the
clitic, ‘there’ -- *Sam* in Hebrew), and the timing of the mouthing of ‘shop’
spans the duration of the host plus clitic together (i.e., of ‘shop there’). In
the vast majority of coalesced forms in our corpus (14 out of 19), the signers
used mouthing, and in every case, they mouthed the Hebrew translation of

\(^{21}\) In ASL compounds consisting of one double-handed sign and one one-
handed member, two different forms may result. The one-handed sign may
come double-handed, or the double-handed sign may drop the
nondominant hand (Sandler 1993c). These forms attest to the high rank of
SYM in the ASL lexicon.
the host word only, and the timing span of this mouthing clearly extended over both the host plus the clitic.\textsuperscript{22}

To sum up, there are two kinds of evidence that coalescence yields a form that resembles a single word. The first kind of evidence shows that the cliticized forms share prosodic characteristics with optimal prosodic words. Although faithfulness to the input would require disyllabicity, the coalesced forms are monosyllabic, the optimal form of the word in sign languages. These cliticized forms do, however, violate the Symmetry Constraint, which is apparently undominated in the lexicon. The second type of evidence shows that the host plus clitic together are a single word in the mind of the signer: only the host word is mouthed, and its span is the whole form, host plus clitic. We now turn to the second cliticization process.

3.2. Dominant Handshape Assimilation

The second process is complementary to the first in terms of its prosodic position. It occurs where the host and clitic are in a prosodically weak position within the phonological phrase, usually at the beginning. In this process, the cliticizing pronoun retains its movement, but is weakened by losing its handshape and assimilating the handshape of its host.\textsuperscript{23} One reason for claiming that this process is cliticization, rather than simply phonological assimilation, is that it occurs between subject pronouns and the following verbs, a typical syntactic relation in cliticization.\textsuperscript{24} Another reason is that systematic assimilation of handshape features has not been reported between words except in the case of pronouns. Handshape assimilation can occur in compounds, however, in both ASL (Sandler 1987, 1989) and in ISL. But in compounds, additional constraints are obeyed. First, orientation assimilates together with handshape (Sandler, ibid., for ASL), and second, handshape assimilations in compounds tend to cooccur with segmental deletions that result in monosyllabicity. In the clitic assimilation, orientation does not assimilate

\textsuperscript{22} The mouthing patterns are a nice example of the way in which borrowed material from a spoken language is reinterpreted, essentially becoming part of the sign language. Thanks to Mark Aronoff for this observation.

\textsuperscript{23} Handshape assimilation between pronouns and verbs is reported to occur in ASL as well (Liddell and Johnson, 1989; Corina and Sagey, 1989; and Wilbur, to appear).

\textsuperscript{24} In our data, all examples were first person pronouns, but this is an artifact of the elicitation sentences, since it is unfelicitous to use personal pronouns other than first person in out of the blue contexts.
together with handshape\textsuperscript{25}, and no segmental deletion within the host occurs either, so that the resulting host plus clitic forms are disyllabic. The effect of the process is to weaken the pronoun by removing the contrast between its handshape and that of its host. The pronoun is usually a proclitic but may also be an enclitic if the basic word order is deviated from, provided host and clitic are not in a prosodically strong position. If the personal pronoun exceptionally follows the verb but in a prosodically strong position, then coalescence may result -- there is one such example in our corpus.

Figure (19) shows the first person subject pronoun in citation form, and figure (20a,b,c) shows the cliticized first person pronoun followed by the verb whose handshape it assimilated -- I READ, from a sentence meaning, ‘I read the story fast’\textsuperscript{26}.

\begin{figure}[h]
\centering
\includegraphics[width=0.2\textwidth]{figure19.png}
\caption{I (citation form)}
\end{figure}

\textsuperscript{25} The fact that orientation assimilates with handshape in compounds motivates a hierarchical representation of these two categories, such that handshape dominates orientation, as shown in figure (3) (Sandler 1987, 1989). Apparently, the features which are hierarchically organized in the lexicon get linearized at some point, allowing handshape to assimilate without orientation. Note that the coalescence process also requires collapse of the feature geometric representation: compare figure (14a) with figure (18). Such linearization is independently motivated in Sandler (1993b).

\textsuperscript{26} The facial expression in figures (20a,b,c) accompanies the meaning, ‘fast’. Facial expression in ISL is argued in Nespor and Sandler (to appear) to correspond to intonation.
In figure (19), the handshape of the pronoun in isolation is an extended index finger. In (20a), the handshape is ‘V’, or the index and middle fingers extended and separated – the same handshape as in the host, READ, shown in (20b,c). In figure (20a), the nondominant hand (h2) is already in its position as place of articulation for the host, READ. This type of spreading of the nondominant hand is analyzed in Nespor and Sandler (1997, to appear) as an external sandhi rule whose domain is the phonological phrase. It is not related to cliticization. Rather, it is the handshape assimilation on the dominant hand that is of interest here.

While the clitic in this process is always a pronoun, the host is occasionally syntactically idiosyncratic, suggesting prosodic restructuring (Nespor and Vogel 1986). An example of this is the sequence TWELVE-I, where the consultant signed the sentence number and, rather than making the usual intonational phrase break, began to sign the sentence as part of the same prosodic constituent as the sentence number. One can easily imagine an equivalent in a spoken language elicitation like this one, where saying the sentence number is very redundant by the time you get to number twelve, and the speaker runs the sentence number into the first intonational phrase of the example.

Like the coalescence process discussed in Section 3.1., which seems to obey a monosyllabicity constraint, this process is also seen as obeying a constraint on the prosodic word, the Selected Fingers constraint, discussed in connection with lexicalized compounds in Section 2.2. If, as suggested there, this constraint holds on the domain of the prosodic word, it offers insight into assimilation under cliticization. In particular, assimilation obeys a constraint on prosodic word form. The cliticized form is disyllabic, however, since two movements are articulated in succession – that of the clitic and that of the host. Thus the monosyllabicity constraint is violated,
indicating that the SFC dominates the Monosyllable Constraint in creating these forms: SF > MONOSYL. The data on fingerspelled borrowing in Section 2.2. indicate that at the lexical level, the ranking of the monosyllabic constraint and the selected finger constraint is the opposite: MONOSYL> SF.

The main points of this analysis are that the words of sign languages obey output constraints on their prosodic structure, and that those constraints are active both lexically and postlexically, although their effects are somewhat different at each level. These theoretical points -- and their relation to more traditional phonological rules -- are discussed in the following section.

4. Constraints and levels.

The effects of cliticization can be described either in terms of phonological processes or in terms of constraints. If the two were merely notational variants of each other, then either would do. But, as the exposition in Section 3 has shown, analyzing the phenomena described here in terms of constraint interaction is more explanatory. Therefore, such an analysis is preferable.

If we look at coalescence as a process as shown in (21), occurring perhaps by association to a template, with some other rules to ensure just the right kind of linearization of features, etc., this would give the correct result. But that type of rule formulation alone would imply that the fact that the output has the optimal monosyllabic form of a prosodic word is a coincidence.

\[(21) \quad \text{LML} + \text{LML} \quad \rightarrow \quad \text{LML}\]

Similarly, if we view handshape assimilation merely as a process like that shown in (22) (as in Sandler 1987, 1989), we miss the generalization that prosodic words optimally have only one handshape. Figure (22) shows a schematic representation of the assimilation in phonological form. The handshape is represented as the features (F) of the Selected Fingers (SF) node.

\[27\] Evaluation of constraints in two stages has also been proposed for spoken languages (Kenstowicz 1994, Booij 1997). However, in those studies, the arguments adduced for this evaluation favor ordering of evaluation (i.e., lexical before postlexical), and not re-ranking.
To account for these phenomena, I have proposed constraints on the prosodic word, such as Monosyllabicity and the SF constraint, and allow them to operate both lexically and postlexically. This is attested in spoken languages, and, according to Booij (1994), is expected in a theory in which rules apply wherever they can. If we take this approach, then we have an explanation for why cliticization should involve just these sorts of processes: at the level of the grammar where morphosyntactic words combine, hosts plus cliticizing pronouns are ‘trying’ to reduce to a single prosodic word. I hasten to note that more constraints and interactions must be motivated in order to ensure precisely the attested surface forms without rules. Given the complexity of the data and the fact that constraints are intended to be universal (at least within a language modality), much more work must be done. However, the Monosyllabicity, Selected Finger, Place, and Symmetry constraints are already well motivated in ASL research, though they have variously been characterized as rules, conditions, or tendencies, and their domains have been proposed to be the ‘sign’ (morphosyntactic word?), the morpheme, or the syllable. The present analysis proposes specific constraints, the prosodic word domain\textsuperscript{28}, and particular interactions which together are hoped to offer the beginning of an explanation of the phenomena under investigation.

5. Unresolved issues

Two different constraint rankings, both postlexical, are required by this analysis of the two types of clitics, a point noted by a reviewer of this article. In coalescence, MONOSYL outranks SF, since the surface forms

\textsuperscript{28} See also Brentari (to appear) for a treatment of the prosodic word in ASL.
involve two different SF specifications on the dominant hand but only one syllable, while in assimilation, SF outranks MONOSYL, since the surface forms are just the opposite. This of course requires some explanation, and, as I have suggested, the explanation is thought to lie in the relative prominence of different positions within larger prosodic constituents. A constraint based explanation would require motivation of phrase-level constraints and their interaction with word-level phenomena. At this point in my understanding, such an analysis would be ad hoc, and I therefore leave it to future research. Another question is why the assimilation clitics retain their syllabic ity although they are in a nonprominent position. I have suggested that the phrase-final pronouns coalesce because they are in a prosodically prominent position but cannot get phrase level stress by themselves. Apparently, there is no need for pronouns to coalesce with the syllable of the host when they are in a prosodically weak position. They remain syllabic, but weakly stressed. Clearly, this important question deserves a more detailed answer, and I leave that to future investigation as well.

As the topic of this volume is the phonological word, it is relevant to address the question of what type of constituent is formed by cliticization: prosodic (or phonological) words, or clitic groups. A distinction between these two was motivated in Nespor and Vogel (1986), but the existence of the clitic group has since been called into question. One might say that the coalescence and assimilation phenomena involve constraints and rankings whose domain is the clitic group, since the precise behavior of these forms does not seem to be occur in any other domain. But, as we have seen, conclusive evidence distinguishing morphosyntactic words, prosodic words, and clitic groups in any sign language is not yet at hand. What the present investigation has shown is this: there are constraints on well-formed words in sign languages which are essentially prosodic in nature (see also Brentari 1990, to appear), and that the cliticized constructions obey some of them. That is, it is not just any constraints that determine the form of these constructions, but rather constraints that make these forms similar in particular ways to well-formed prosodic words of the language.

---

29 This question was raised by Laura Downing (p.c.).
6. Conclusion: What is really universal?

Words are often pronounced differently in connected speech than they are in isolation (Kaisse 1985). These differences are constrained by both syntactic and phonological factors. Syntactic domains themselves have prosodic properties associated with them -- primarily, rhythmic and intonational patterns. All of these elements interact with each other systematically in language in such a way as to enhance communication.

This study has presented an example of such interaction in a language in a different physical modality. The constraints on the phonological forms of the words of Israeli Sign Language were shown to have an effect when words are joined together in connected signing as well. In particular, cliticization is seen as imperfect prosodic word formation in connected signing. In the course of the investigation, evidence was provided to show that a sign language bears interesting similarities to spoken languages at the interface between phonology and syntax. As always, though, research on sign language presents interesting challenges too. In particular, it raises the question: What is universal in phonology?

Languages in both modalities have duality of patterning, i.e., a phonological level that is distinct from the meaningful level. Both have complex words that are formed in a systematic and conventional way. Phonological form often changes in response to morphological operations in both modalities. The present study shows that the words of sign language, like those of spoken language, conform to certain well-formedness constraints, and that some of the same constraints are also active when words combine in sentences, resulting in a change of phonological form vis a vis the input at this level as well.

Yet some things are different. The constraints of OT, for example, are intended to belong to a universal set, perhaps available to the child at birth. But clearly, sign language constraints are different. Where spoken languages have modality-specific constraints on syllable structure such as NO CODA or feature spreading constraints such as PAL, sign languages have constraints such as MONOSYL and SYM, and a different set of features. These differences are not trivial, as they show that some constraints are universal only within a particular modality, raising the following fundamental questions: Are children genetically equipped with the full bag of constraints for both modalities? Or is it more
reasonable to hypothesize that the constraints arise through experience in
response to production and perception pressures of the modality in which the
language is transmitted?

One might expect more general families of constraints such as ALIGN or OCP
to bridge the two modalities, but whether this is the case, and whether such
families really behave the same way in the two modalities, are empirical
questions. When we can answer questions such as these, we will understand a
good deal more about human language than we would if these questions never
arose. It is only through comparative investigations of the two natural human
language modalities, spoken and signed, that these important questions can be
raised.

References
University Press
Aronoff, Mark, Irit Meir and Wendy Sandler. In preparation. Words in Space:
Morphological Characteristics of Sign Languages.
Spring: Linstok Press
Booij, Geert. 1994. Lexical phonology: a review. *Lingua e Stile a. XXIX.*
525-554.
Booij, Geert. 1997. Non-derivational and lexical phonology. In *Derivations and
Brentari, Diane. 1993. Establishing a sonority hierarchy in American Sign
281-308.
Cambridge, MA: MIT Press.
Brentari, Diane and John Goldsmith. 1993. Secondary licensing and the
nondominant hand in ASL phonology. in G. Coulter, ed. 19-41.
Corina, David. 1990a. Reassessing the role of sonority in syllable structure:
Evidence from a visual-gestural language. In *ParaSession on the Syllable*


Hulst, Harry van der. 1996. On the other hand. Lingua 98:1-3. 121-143


Sandler, Wendy. in preparation. *Optimal Signs.*
Sandler, Wendy and Diane Lillo-Martin. in preparation *Sign Language and Language Universals*