Tommi Jantunen

How long is the sign?

Abstract: This paper deals with the relative empirical length of signs in sign languages and provides evidence for the view that they are actually longer units than has hitherto been recognized. The evidence is presented from two perspectives: those of sign articulation and sign recognition. Concerning sign articulation, it is suggested that signs are longer units than is currently assumed because most of the structural features of signs are in fact already present before the currently accepted beginnings of signs and they continue after signs' generally accepted endings. Concerning sign recognition, the longer view of the sign is proposed on the grounds that the recognition point of signs is typically located before their alleged beginning, and because signs (as currently understood) can also be recognized on the basis of parts of their subsequent transitions only. The nature of the longer sign is discussed together with some more general consequences for sign language research of the revision of our view of what a sign might be.

Keywords: sign, linear sign, length of sign, sign articulation, sign recognition, phonetics, sign language

DOI 10.1515/ling-2014-0032

1 Introduction

Sign language linguists have been investigating the nature of the *sign* ever since 1960, when William C. Stokoe published the first structuralist study on American Sign Language. In his study, Stokoe (1960) showed how signs are composed of smaller *simultaneously* realized sub-components of *handshape*, *place of articulation*, and *movement*. Because these components are capable of forming minimal pairs, Stokoe analyzed them as the sign language counterparts of spoken language phonemes. After Stokoe's early work, other types of phonemic elements have also been identified in the sign. For example, since Battison (1978), hand *orientation* has been included in the list of the sign's structural sub-components, and today researchers also give this status to the *non-manual element* which, in

Tommi Jantunen: Finnish Sign Language, Sign Language Centre, Department of Languages, P.O. Box 35 (F), FI-40014 University of Jyväskylä, Finland. E-mail: tommi.j.jantunen@jyu.fi

discussion of sign structure, refers primarily to the movements and positions of the mouth (e.g., Sutton-Spence and Woll 1999; Johnston and Schembri 2007).

In the course of research, more substantial revisions have also been introduced to Stokoe's original model and its successors (e.g., Brennan 1980; Rissanen 1985). Of these, one of the most important has perhaps been the adoption of the notion of *structural sequentiality* in the study of signs. In Stokoe's model, the temporal organization of handshapes, articulation places and movements (and orientations and non-manual elements) within signs and sequences of signs were not seen as an important object of study. However, in 1984 Scott K. Liddell published a study in which he claimed that sequentiality is a phonologically important feature of signs and that, in fact, signs may be analyzed into strings of Hold and Movement *segments*, also capable of expressing phonological contrast. Consequently, since Liddell (1984), most work on sign structure has come to acknowledge the importance of sequentiality – and some types of segments – in the analysis of signs (e.g., Liddell and Johnson 1989; Sandler 1989; Perlmutter 1992; Brentari 1998; Johnson and Liddell 2011).

Although the research so far carried out into the nature of the sign has undeniably advanced our knowledge about the unit, there still remain many aspects of signs that are largely uncharted. One such aspect is the relative empirical length of a sign or, in other words, the question of where in the sign stream the *linear*, temporally delimited sign begins and where it ends. Apart from a few early attempts (e.g., Green 1984), this question has not been directly tackled in the modern discipline of sign language linguistics. Instead, there is an unspoken consensus that the linear sign is a movement period during which the hand either travels between two spatial placements (these placements, between which the hand may perform a variety of sub-movements [e.g., handshape and/or orientation changes], are ultimately identified by a change in the direction of the movement of the hand), or executes some type of hand-internal movement (e.g., finger wiggling) without performing any noticeable global motion in space; some signs may even be periods of stasis. The consensus is, further, that the sign may be one or two-handed, and if the latter, it is the activity of the dominant hand that is used as the point of reference for the linear delimitation of the sign. This conception of the length of signs was already present in Stokoe (1960) and, as evidenced by, for example, the way in which signs are currently identified and annotated in corpora (e.g., Crasborn and Zwitserlood 2008; Wallin et al. 2010), it forms the normative ontological and methodological basis for all modern work in sign linguistics, regardless of the theoretical framework or the level of analysis.

However, the past few decades have seen an accumulation of evidence that suggests that the current conception of the linear sign may not be accurate enough. More specifically, a considerable number of phonetic findings seem to

indicate that the borders of the linear sign actually lie outside the empirical domain that is currently considered to be the sign and that the linear sign – as a normative class notion introduced above – is a longer unit than is currently recognized. One main type of evidence supporting this claim comes from study of the articulation of signs (i.e., from articulatory phonetics): this work has shown, for example, that several aspects of sign structure (e.g., handshape, or nonmanual elements, or the other hand) are typically present well before the generally accepted beginning of a sign, that is, prior to the production of the location in which the hand starts to articulate the movement typically associated with the sign (e.g., Jantunen 2010, 2011; Johnson and Liddell 2011). Another type of evidence comes from research done on sign recognition (which, because of its focus on psychological issues related to sign processing, may be equated with auditory phonetics): this work has shown, for example, that those sequences of sign stream corresponding to our present view of the sign are, in fact, paradoxically not a necessary requirement for sign recognition; that is, signs are in many cases already recognized before their production is even supposed to have begun (e.g., Grosjean 1982; ten Holt et al. 2009).

This paper sets out to discuss the length of the linear sign. More specifically, the paper first describes in detail the present mainstream view of the sign (Section 2). The main contention here is that the current mainstream conception of the length of the sign is based on very early work carried out to describe the movement of the sign. After this, the paper goes on to present evidence suggesting that the sign is linearly – that is, in terms of its starting and ending points (regardless of its inner structure, complexity, or articulatory context) – a longer unit than the current mainstream view acknowledges (Section 3). The evidence is approached from two perspectives: those of sign articulation (3.1) and recognition (3.2). Finally, the paper outlines briefly the nature of the longer sign that emerges on the basis of the evidence presented here and discusses some of the possible consequences for sign language linguistics in general of such a revision of the sign (Section 4).

In this paper, the linear sign is understood as a concept in *phonetics* and hence, by definition, the main discussion of the subject is within that framework (both articulatory phonetics and auditory phonetics; Section 4 even discusses briefly acoustic phonetics). However, most major work on the structure of the sign, reviewed below in Section 2, has been *phonological* in nature. As phonological organization is more abstract than phonetic organization (e.g., Ohala 1997), the question arises of whether phonologically based discussion about the length of the sign is at all relevant in phonetically based discussion. Indeed, given the fact that sign language phonologists have not typically explicated the link between phonology and phonetics (see Crasborn 2012), making any *formal* claims

about the phonology-phonetics interface is unwise. However, it must be emphasized that such an endeavor is not the purpose of the present paper. Rather, the goal in presenting phonological work (in Section 2.1) is simply to illustrate the underlying, shared, deeper conception about the sign and its presumed length in relative terms.

Nevertheless, concerning the link between phonetics and phonology, the present paper operates on several assumptions that have to be explicated. First, it is assumed throughout the paper that phonology *is linked* to phonetics, although the formal nature of this link may vary between theoretical frameworks and models. Second, as will be shown, for example by the quotations presented in Section 2, this link is assumed to be in many cases even a fairly direct one. In practice, this means that if any sign is to be described phonologically, for instance, as having a straight movement, the phonetic manifestation of the sign is then presumed to be, in the prototypical case, a relatively straightforward movement phase in which the hand proceeds from one describable placement to another, first gaining speed and then slowing down, both describable placements identifiable by a change in the direction of the movement of the hand. Finally, concerning especially the more general discussion about the possible phonological consequences in Section 4, it is assumed that phonology *should also be based* on phonetic facts. I acknowledge that this is by no means a necessary requirement for phonological work, which is abstract by definition (e.g., Ohala 1997). However, I am convinced that if the ultimate goal is to investigate language that people actually use, such an approach must be followed.

2 How long do we think the sign is?

This section outlines the mainstream view that sign language linguists currently hold concerning the length of the linear sign. Section 2.1 first discusses the sign from the perspective of phonology and the most influential phonological models of sign structure. Section 2.2 changes the perspective to that of phonetics.

2.1 Phonological perspective

In general, the question of where in the sign stream the linear sign begins and where it ends has attracted little research attention. At the beginning of the modern era of sign language research and phonology, the question was not considered to be important, as Stokoe (1960) and his followers defined the sign as a structurally *simultaneous* unit:

> The sign-morpheme, however, unlike the word, is seen to be not sequentially but simultaneously produced. Analysis of the sign cannot be segmented in time order but must be aspectual. The aspects of the sign which appear to have the same order of priority and importance as the segmental phonemes of speech are the aspects of [hand] configuration, position or location, and motion. (Stokoe 1960: 39–40.)

For Stokoe, in addition to his conviction that the sign was not a temporally organized unit, any question about the empirical borders of the sign was also rendered unimportant by the methodological tradition of structural linguistics, which was primarily interested in identifying the abstract *emic* components of a language. Within this tradition, there was no reason to ask such questions as where a morpheme begins, because the methodology of the tradition was devised to describe only the internal components of the morpheme. In practice, this was achieved through the use of descriptive taxonomic symbols to label the phonemic/morphemic and allophonic/allomorphic elements of a language.

Although Stokoe and his followers did not consider the sign as a linear unit, it is reasonable to assume that, had they done so, they would have taken the length of the sign to be the same as the length of its movement(s), which they classified either with directional labels such as *up*, *right*, or *toward*, or – in the case of smaller movements produced from the wrist or finger joints – as *supinated* or *open*. When applied, for example, to the current isolated, one-handed monomorphemic Finnish Sign Language sign BLACK, shown in Figure 1, this ideology still identifies the linear form generally associated with the sign. In Stokoe's model, BLACK has a (straight) movement toward the body, and it begins at the location in which the hand first starts to move toward the signer; in Figure 1, this

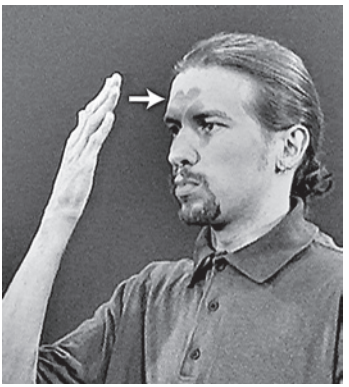

Fig. 1: Finnish Sign Language sign BLACK as depicted in the Basic Dictionary of Finnish Sign Language (Malm 1998).

is the position in which the hand is depicted. The end of the sign is the location in which the hand stops articulating this movement. In the case of BLACK, this location is on the forehead, as shown in Figure 1 by the arrow describing the direction of the movement.

In BLACK and in other isolated signs, the hand needs to be first raised to the beginning location and, after the production of the movement, lowered from the ending location back to the resting position. These very initial and final movements of the hand (not shown in Figure 1) were not treated as parts of signs by Stokoe and, indeed, in the modern work of sign language phonology they are still classified as *transitional movements*, or *transitions*, and excluded from linguistic analysis (for more on transitions, see Jantunen 2013).

The need to be more specific about the nature of the borders of signs became more important after Liddell (1984) and the introduction of the idea that, contrary to what Stokoe had proposed, signs also have a sequential structure. In Liddell's original Movement and Hold model (MH model), the question of the length of the sign was answered through the notions of Movement (M) and Hold (H) segments. Ms were defined as periods of time when the articulator (typically, the hand) was in motion; Hs were the moments when all the postural features of the articulator remained unchanged. In the MH model, signs were seen as being composed of these segments and typical signs were described as M, H, MH, or HMH type of sequences. The segmental structure of signs was first decided on the basis of the activity of the dominant hand and this was then imposed on the description of the non-dominant hand; in practice, the segmental structure of the non-dominant hand was different from that of the dominant hand only in cases where the non-dominant hand acted as a static place of articulation. In the case of the sign BLACK in Figure 1, the MH model representation of its segmental structure was MH, capturing the fact that, in the production of the sign, the hand stays in contact with the forehead for a short while.

The Ms and Hs in Liddell's (1984) original MH model were introduced as *a priori* notions whose exact status as phonetic or phonological units was not, and still is not, very clear. However, what is known on the basis of Liddell's text is that Ms and Hs were connected *post hoc* to empirical video data so that Ms were identified with periods of time when the contour of the articulator appeared blurred in the video frames; Hs were identified with the times when the contour of the articulator was clear. This idea is expressed, for example, in the following:

> In an effort to get a better understanding of the importance of movements and holds for signs in general, videotapes have been examined [...]. When looking at any given field, it is possible to see whether the hand is moving along a path: a blurry image of the hand in an otherwise clear picture indicates that the hand is moving. (Liddell 1984: 377.)

All in all, the order of presentation in Liddell (1984) suggests that original Ms and Hs should be taken not only as phonological (i.e., abstract) units, but also as units which relate to phonetic (i.e., concrete) behaviors in a relatively straightforward way (i.e., Ms represented visible movement periods and Hs observable static periods in the activity of the hand).

The original MH model (Liddell 1984) has been a reference point for all subsequent major phonological models of sign language phonology: Liddell and Johnson (1989), Sandler (1989), Perlmutter (1992), and Brentari (1998). In practice, this has meant that the view of the length of signs – their beginnings and endings – advocated by the MH model has also been accepted in these other models. This is equally true for models that operate with MH-style classical dynamic and static segments as well as for models in which segments are understood in an even more abstract manner. An example of the first group is the phonological Hold-Movement model devised by Liddell himself with his colleague Johnson (Liddell and Johnson 1989). Another example is the Hand Tier model introduced by Sandler (1989), in which the prototypical sign is represented as a sequence of Location-Movement-Location segments. The relationship of phonological Location (L) and Movement (M) segments to the actual phonetic behavior of the hand in the production of the American Sign Language sign INTELLIGENT is described by Sandler (1993) as follows:

> Consider the sign INTELLIGENT [...]. This sign is characterized by the canonical LML form. The first location is in contact with the side of the forehead, ipsilateral to the signing hand [i.e., on the same side of the signing hand]. This is followed by a straight movement, and a second location, a short distance in front of the first. As is typical of the vast majority of monomorphemic signs, there is a single hand configuration throughout [i.e., the upward pointing index finger extended from the fist]. (Sandler 1993: 245.)

An example of a model in which segments are more abstract units is Brentari's (1998) Prosodic model. In this model, the segmental structure of the prototypical sign includes two x-slots that are linked together with a suprasegmentally realized movement component. Figure 2 demonstrates the Prosodic model concept of the sign by showing schematically how it represents the isolated sign BLACK (Figure 1).

In the Prosodic model (Brentari 1998), signs are considered to correspond to sequences of sign stream that can be described phonologically by using a group of hierarchically organized inherent (or static) and prosodic (or dynamic) features, the latter (PF) occurring prototypically between two timing units, x-slots (inherent features, IF, are represented collectively with a triangle in Figure 2). Of the two types of features, only the prosodic ones specifying the movement of the sign are realized sequentially (the path feature meaning 'directed toward a plane'

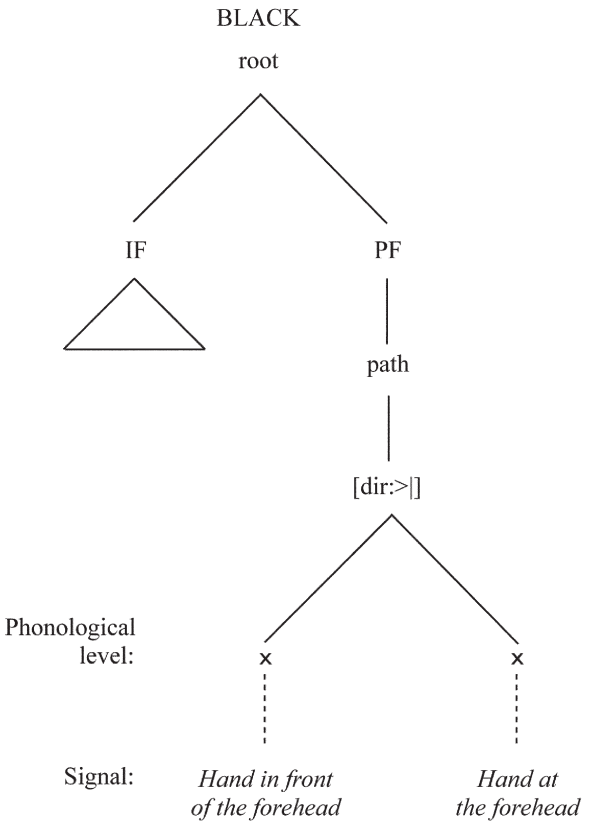

Fig. 2: Prosodic model (Brentari 1998) representation of the Finnish Sign Language sign BLACK (see Figure 1).

in Figure 2) and hence only they define the sign linearly. Like the original MH model (Liddell 1984) and other segmental phonological models (e.g., Liddell and Johnson 1989; Sandler 1989; Perlmutter 1992), so the Prosodic model treats movements outside sign borders as transitional and excludes them from phonological and linguistic analysis.

The representation of the non-dominant hand of two-handed signs is even more reliant on the representation of the dominant hand in the Prosodic model than in segmental models. This is because in the Prosodic model the non-dominant hand is treated as a dependent structural unit of the dominant hand. Moreover, because the non-dominant hand is represented only with inherent features, it is assumed to be fully constrained by the sequentiality and linearity manifested by the dominant hand.

It must be emphasized that the link between phonetics (the concrete signal) and phonology (the abstract representation) is not fully explicated in any of the phonological models discussed in the present section. In other words, it is not fully known how to associate the abstract phonological representation of the models with the properties of the actual signed signal, or vice versa. However, as evidenced by, for example, the cases referred to above – which are fully in line with the underlying thesis of the present paper (see Section 1) – such an association can undoubtedly be made, and indeed quite straightforwardly (cf. Crasborn 2012: 7, who speculates that this "straightforwardness" may be due to the fact that we can *see* the articulators involved in the production of the phonological form). That this is the case also in the more abstract Prosodic model is evident, for example, in the way the model treats the timing of the handshape and orientation change features and path movement features, as well as in the fact that the model is built to capture certain inherently phonetic phenomena (e.g., the weakening and strengthening phenomena of movements and the phrase-final lengthening of signs, see Brentari 1998; see also Sandler 1989, Perlmutter 1992, and Crasborn 2001).

2.2 Phonetic perspective

As I have noted above, Liddell's (1984) work really marked the beginning of the investigation of signs as temporally organized units. However, surprisingly, it left the actual criteria for defining the borders of the sign rather vague. In the context of phonetic models, these criteria have been examined by Liddell and Johnson only recently, over 25 years after the publication of the original MH model. In their second paper in a series of a total of eight papers in which they describe their new phonetic notation system (Johnson and Liddell 2011), historically based on the MH model (Liddell 1984) and its clearly phonological successor the Hold-Movement model (Liddell and Johnson 1989), they commit themselves explicitly on the question of how to determine the borders – and consequently, the length – of the American Sign Language sign CHICAGO (see Figure 3):

> [Concerning the identification of the parts of the sign stream that are or are not part of the sign CHICAGO:] We employed three criteria for these decisions. First, when we look at other instances of CHICAGO, we see that they all tend to share a similar starting place [...]. Directly before this placement, however, the activities of the hand vary as a result of the preceding context. [...] Because these transitional activities vary from utterance to utterance, we do not treat them as a part of the sign itself. [...] Second, note [...] that the hand moves upward from its resting position on a direct and straight path that ends with [...] the beginning of the sign. After [this] the hand moves along a different, lateral path. This change

Fig. 3: American Sign Language sign CHICAGO.

> in direction also provides evidence for the beginning of the sign. Third, when we ask native signers to show us how CHICAGO begins, they always place their hand in roughly the position shown in frame 1 [i.e., in the juncture between the two movements]. (Johnson and Liddell 2011: 412.)

Johnson and Liddell (2011) describe the use of three criteria to determine the beginnings of signs. The first criterion requires one to compare different instances of the sign and determine the invariant starting place on the basis of this comparison; periods during which in contextual and continuous signing the hand is approaching this starting place from different directions are classified as transitions. The second criterion, supporting the first one, asks one to associate the beginning of the sign with the location at which the direction of the movement of the hand changes; this is the same location as marks the operational beginning of movement features (e.g., *straight*) in phonological models (see Section 2.1). The final criterion involves consulting the intuition of native signers, who are said to be able to roughly approximate the beginning of the sign with the convergence of the two movements. It is worth noting that no actual description is provided by Johnson and Liddell for the process of identifying the ends of signs.

The criteria Johnson and Liddell (2011) describe resemble closely the phonetic sign identification criteria that have recently emerged in the new field of sign language corpus linguistics. Building a representative sign language corpus requires one to first collect and then annotate a large amount of sign language video material. This in turn calls for the linear identification of signs, or corresponding video sequences, from the video stream. For this task several guidelines have been established (e.g., Crasborn and Zwitserlood 2008; Johnston 2009; Wallin et al. 2010). The guidelines differ slightly from each other but they all have one thing in common: they all take the bipolar sequential movement phase of the

dominant hand as the starting point for the sign identification. This is demonstrated, for example, in the following excerpt from the annotation guide for the Corpus NGT (the Sign Language of the Netherlands):

> [A sign starts] at the first frame in which the hand starts to move away from the initial location of the sign to the final location of the sign; or (in case the hand does not move through space) [at] the first frame in which the handshape starts to change [...]; or (in case the hand does not move through space **and** the handshape does not change) [at] the first frame in which the orientation of the hand starts to change. [A sign ends] at the first frame in which the handshape starts to change after the sign was finished; or at the first frame in which the hand starts to move away from the final location of the sign. (Crasborn and Zwitserlood 2008: 6; original boldface.)

Sign identification criteria such as the one described above may be applied to both isolated signs and to continuous signing, in both of which they serve to separate signs from the surrounding transitions. When applied to continuous signing, the criteria result in a division of the sign stream into longer sequences of linguistic signs and non-linguistic transition phases, as presented in Figure 4. Within this type of approach, the average duration of signs has been calculated to be roughly a third longer than that of transitions, and the duration of signs has also been claimed to exhibit more variation than that of transitions (Jantunen

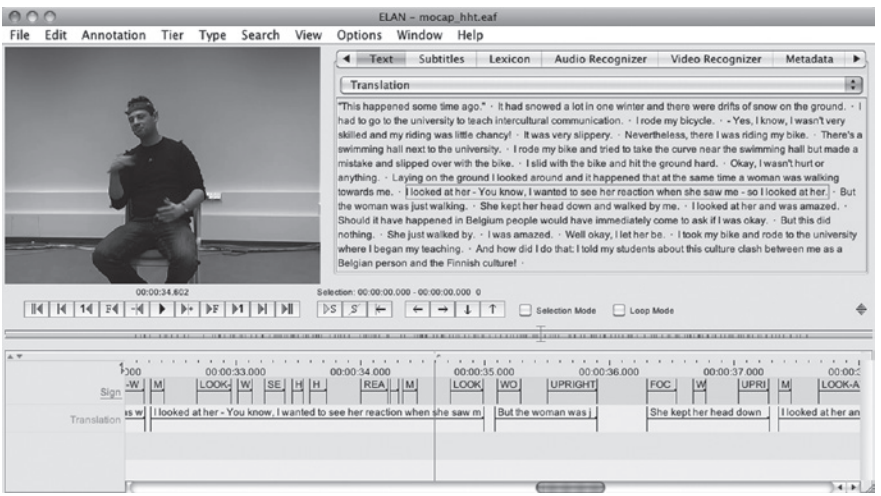

Fig. 4: A screenshot from ELAN (The Max Planck Institute of Psycholinguistics, Nijmegen, The Netherlands; http://tla.mpi.nl/tools/tla-tools/elan/) showing sign annotations and gaps between the annotations (the upper tier). Data from Jantunen (2013).

2013). An interesting implication of all this is that a relatively large proportion of actual signing is in fact considered to be linguistically redundant activity. This is not the way a speech signal is seen to be composed.

The dominant hand dominates the corpus-related annotation work (e.g., Wallin et al. 2010). Many corpora nevertheless distinguish between the dominant and non-dominant hand, that is, annotate them as different units (i.e., on different tiers). I consider this to be an important and welcome methodological trait as the two hands are capable of performing very different articulatory activities in signing. Furthermore, the separation of the two hands in corpus annotation often also indicates that the temporal relations within articulators are not necessarily a straightforward matter (for more, see Section 3.1).

The sign identification criteria included in corpus annotation guides are currently the most accurate phonetic criteria available for deciding the beginnings and ends of signs. In general, their emergence reflects a shift in sign language research from intuition-based research toward the more extensive collection and exploitation of data. However, it is claimed in this article that even these most recent criteria – used today in corpora-based research in the fields of morphology and even syntax – are based on an abstract conception of the linear sign that ultimately draws on early phonological work on sign languages, as has been documented in this section. How accurate this phonologically grounded mainstream view of the linear sign really is will be discussed in the following sections.

3 Phonetic evidence that the sign may be a longer unit than is assumed by the present mainstream view

As I have said in the previous section, modern research into sign languages operates on the phonologically based assumption that the linear phonetic domain of the prototypical sign is a movement phase in which the (dominant) hand either proceeds (with or without hand-internal movement[s]) from one describable placement to another describable placement, both identifiable by a change in the direction of the movement of the hand, or executes a hand-internal movement in one place. In some cases, the hand may also remain immobile. This section now looks at the phonetic evidence that in the long run is likely to be recognized as showing that signs are longer units than is generally currently recognized. Evidence is presented from two perspectives: those of articulatory phonetics (3.1), and sign recognition (3.2).

3.1 The perspective of articulatory phonetics

Let us begin by briefly considering the production of a spoken word. Where does the phonetic form of a spoken word begin and where does it end in terms of articulation? A fairly common sense answer to this question is that the beginning of a word (or any phonetic unit) is determined by the first occurrence of any articulatory feature that is typically associated with the unit and, similarly, that the end of the word (or any unit) is determined by the last occurrence of the last of such features. For isolated words, determining beginnings and endings in this way is relatively straightforward. For words that occur in continuous speech, on the other hand, the case is a bit more complex: due to coarticulation, articulatory features are mixed and spread so that the phonetic borders of words are never exactly clear-cut or sharp, but fuzzy, words being fused together into a continuous stream. However, as indicated by the success of phonetic research into spoken language (see Hardcastle and Laver 1997 and papers therein), this is not an insurmountable problem for successful speech-to-word segmentation.

Do the currently assumed beginnings and endings of the phonetic forms of signs correspond to the moments of the first and last occurrences of articulatory features forming the signs? Surprisingly, on the basis of a number of phonetic studies, it seems that the answer to this question is very often no. For example, in their analysis of the phonetic execution of different parameters in a sign, Emmorey and Corina (1990: 1237) state that "[m]uch phonetic information appears to be available quite early". Similarly, in their phonetic appearance-based description of movement phases in signs, Kita et al. (1998: 28) conclude that these units always begin with a phase during which "the hand shape and the orientation of the hand are set to the starting values of the expressive phase [the stroke, i.e., the phase typically identified as the sign]", and that they also always end with a phase during which these values are retracted from the end values of this phase. Ojala (2011), in turn, has shown how there are even finger-specific differences in the preparation speed both to and from the target handshape, the index finger being the one that controls the production of the sign and its phases the most.

Indeed, phonetic work on sign articulation has indicated that lots of information on especially handshapes and orientations is typically present before the generally accepted beginnings of signs, that is, prior to the production of the location in which the hand changes direction and starts to articulate the movement typically associated with or defining the sign. Moreover, these studies have also shown that articulatory information on these components also remains accessible well after the production of the unit. The current conception of the linear sign seems to cover only a part of the emerging phonetic domain and, if we accept the meta-methodological rationale that a phonetic form begins from the first

occurrence of any articulatory feature associated with the form and ends with the last occurrence of any such features – as is the case with spoken words – then signs must be reckoned to be considerably longer units. To illustrate that this increase in length is indeed considerable, let us consider the example in Figure 5.

The example in Figure 5 is a frame-by-frame (25 fps) representation of the production of the isolated one-handed monomorphemic, yet phonologically complex (i.e., the sign includes both a path and a local movement of the dominant hand), Finnish Sign Language sign TO-SEND captured from the on-line dictionary of Finnish Sign Language, *Suvi*. By following the mainstream sign identification criteria used, for example, by Johnson and Liddell (2011) and corpus annotators (e.g., Crasborn and Zwitserlood 2008; Wallin et al. 2010) (see Section 2.2), we are able to say that the phonetic domain of the sign, according to the mainstream view, corresponds to the frame interval 10–17. However, frame number 10 is not the first frame in which the articulatory features of the handshape and orientation comprising the sign are present for the first time, nor is frame number 17 the last frame in which these features can be detected. As the figure shows, the initial handshape of the sign TO-SEND is already fully formed in frame 5 and the configuration of the final handshape of the sign can still be detected from frame number 20. Likewise, the target orientation of the hand is already reached in frame number 7 and traces of it are still clearly detectable in frame number 19. Consequently, if the sign TO-SEND is segmented from the video stream by taking into account the first and last occurrences of articulatory features comprising it, then it is actually roughly twice as long (in frames) as the current mainstream view assumes.

The case is similar with all types of isolated signs in *Suvi* (i.e., one and two-handed, mono and polymorphemic, those with simple or complex movement etc.). In a recent computer-vision based study into automatic sign identification by Viitaniemi, et al. (2014), the first occurrence of the first articulatory feature in *Suvi*'s signs (n = 1212) was estimated to be, on average, 3.9 frames earlier than the mainstream view assumed. Comparably, the last occurrence of the last of such features was estimated to be on average 1.9 frames later than the mainstream conception of the end of the sign took it to be. The case appears to be no different with signs produced in context: the articulatory features of handshape and orientation are also present in connected signs both before their alleged beginnings and after their alleged ends. This is demonstrated in Figure 6, adapted from Jantunen (2011).

The example in Figure 6 represents, again, frame-by-frame (25 fps), the production of the semantically complex Finnish Sign Language sign with the meaning 'the house is at place *x*' (Jantunen 2011). The sign – produced, following the

Fig. 5: A frame-by-frame (25 fps) representation of the production of the isolated Finnish Sign Language sign TO-SEND captured from *Suvi*.

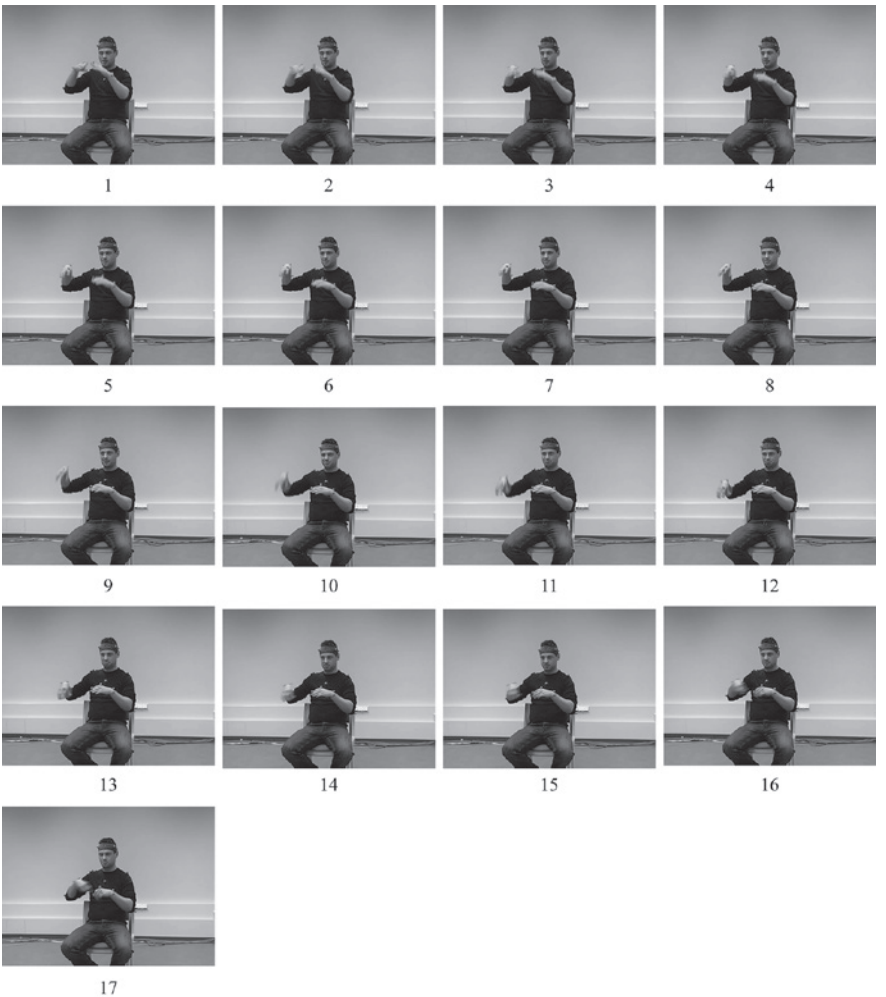

Fig. 6: The frame-by-frame (25 fps) representation of the semantically complex Finnish Sign Language sign with the meaning 'the house is at place *x*' occurring in a stream of signing (Jantunen 2011).

standard way of description, with a palm down claw-like downward moving hand – is extracted from continuous signing in which the signer is telling a story that in this particular moment is describing the topographical location of a swimming hall. The sign is preceded by the compound sign SWIM^HALL – the remnants of which are still visible in the form of a *fragment buoy* in the non-dominant hand – and followed by the first person index finger pointing sign ME; in grammatical

terms, the sign is the verbal predicate of an intransitive existential sentence 'There is a swimming hall at the area.'

Again, according to the present mainstream view, the beginning of the sign is equated with frame number 7, which shows the placement of the hand from which the hand first starts to move downwards. The generally accepted end of the sign, in turn, is shown in frame number 12, which shows the location of the hand after the downward movement is completed. However, as seen from the frames preceding the frame in which the sign is assumed to begin (7), the production of the handshape and orientation of the sign have already begun approximately four frames earlier; the orientation of the hand corresponds to the orientation of the sign already in frame 3 or 4, and the finger configuration of the sign is achieved in frame 4 or 5. Similarly, the handshape and the orientation characteristics of the hand in the sign are continued several frames after frame 12, in which the sign is assumed to have ended; they start to resemble – in a continuous fashion – the orientation and handshape of the following sign ME only after frame 15.

In the example in Figure 6, the activity of the non-dominant hand in the production of the two-handed compound sign SWIM^HALL is very significant in any discussion concerning sign boundaries and the length of signs. If we accept that signs – just like words – begin and end with the first and last detectable articulatory feature, then the sign SWIM^HALL actually ends only near or after frames 16 or 17, when the orientation of the dominant hand begins to relax. In other words, the sign SWIM^HALL lasts throughout the whole articulation of the sign 'the house is at place *x*'. This is a coarticulatory effect caused by the existence of two separate manual articulators capable of functioning independently. There is no strict equivalent of this phenomenon to be found in spoken languages (apart from the fact that it is an example of coarticulation that melts units together) but in sign language research the phenomenon has already been documented, as the following quotation from the annotation guide for Corpus NGT demonstrates:

> In two-handed signs the hands do not always move in exactly the same way. Often one hand stays in a particular position after the sign has ended, while the other hand goes on signing the next sign. Or one hand starts to move or change slightly before the other hand does. (Crasborn and Zwitserlood 2008: 6.)

Articulatory information on handshapes and orientations (of both hands) is not the only type of phonetic information that is available relatively early and late as far as the production of signs understood in the mainstream fashion is concerned. Jantunen (2010) has documented that mouth movements and positions also begin before the presently assumed beginnings of signs and continue after their presently assumed endings. In a study in which he investigated the

durations of signs identified both by the mainstream corpus annotation method and by a method in which the beginnings of signs were identified with the moments the mouthings and mouth gestures of the signs began and ended (a method he called the *non-manual method*), he found that signs identified from short signed sentences by the mainstream method were significantly ($p = 0.0001$) shorter (mean duration 5.9 frames; $SD = 1.8$) than those identified using the non-manual method (mean duration 8.4 frames; $SD = 3.2$). The motivation for Jantunen's use of the non-manual method was that the language he was investigating, Finnish Sign Language, layers signs constantly either with Finnish derivative word-shapes (*mouthings*) or sign language specific *mouth gestures* (see Rainò 2002 for more discussion on these) and, moreover, uses these mouth activities to distinguish the meanings of signs.

In Jantunen's (2010) study, the sequences identified as combinations of a distinct lexical sign and a pointing sign by the mainstream method were counted as only one sign by the non-manual method. This was caused by the regular spreading of the mouth features of the lexical element over the following pointing element: for example, in the production of the sequence CANNOT and pointing, the Finnish mouthing [eei.vo] originating from the Finnish words *ei voi* 'cannot' (lit. 'no can') accompanying the first element was stretched over the following pointing in such a way that the first syllable of the mouthing was associated with the element CANNOT and the second syllable with the pointing. Overall, given that the non-manual element of mouth movement and posture is a structurally obligatory part of Finnish Sign Language (Rainò 2002), this suggests that the two morpheme-like units form, in fact, a single phonetic sign. A similar claim has been put forward, for example, by Zeshan (2002).

Interestingly, the fact that articulatory features of the handshape, orientation and non-manual elements are present well before and after the alleged beginning and ending of the sign has been noted in studies that have explicitly supported the mainstream view of the length of a linear sign. This is demonstrated, in terms of the beginning of the sign produced in isolation, by the following quotation from Johnson and Liddell (2011), in which they discuss the temporal alignment of different structural components – hand configuration (HC; cf. handshape), placement of the hand (PL; cf. place of articulation), facing of the hand (FA; cf. orientation), and non-manual elements (NM) – in the American Sign Language sign CHICAGO:

> Also, note that the HC, the FA, and the NM configuration are established before the hand arrives at the first PL of the sign. Thus, it is not the case that all of the components are established at exactly the same moment and that they all change at exactly the same time. (Johnson and Liddell 2011: 415.)

Regardless of these observations, the beginning of the sign CHICAGO is identified by Johnson and Liddel (2011) with the location in which the direction of the global motion of the hand is changed. The reasoning they offer for this has been explained in Section 2.2 and it rests on the claims that the location in which the hand changes direction is the first invariant feature of the sign and that deaf native signers identify this location intuitively with the beginning of the sign. However, on the basis of our common ability to make phonetic observations and on several phonetic studies, the claims would seem to be not entirely accurate. Concerning the first claim and isolated signs, it is not the case that the route that is labelled *transitional beginning* to the sign actually varies very much. In fact, one might argue that it is a very stable part of the articulation of isolated signs because it remains roughly similar from utterance to utterance, as signers always have to raise their hand with a similar global movement to the point where the movement then changes its direction for the first time. Moreover, as has been demonstrated in this section, this transitional beginning also contains important articulatory information about the sign. It is only in the context of contextual signing that the reasoning of Johnson and Liddell makes sense. However, this context adds another type of difficulty to the claim, namely that caused by coarticulation: several studies (e.g., Mauk 2003; Ojala 2011) have shown that locations Johnson and Liddell treat as invariant actually vary considerably in continuous signing; that is, signs and signing may be lowered or raised for a variety of reasons, or be produced in grammatically marked locations. Consequently, it seems that it is not the case that the claim about the invariancy of the first location made by Johnson and Liddell actually holds true. Concerning native intuitions, on the other hand, the next section will show that, according to more detailed studies, for example, on American Sign Language, native signers actually associate the beginnings of signs before the location Johnson and Liddell, among others, assume to be the beginning of the sign.

To summarize, this section has shown how most of the structural features of the sign – handshape, orientation, and non-manual elements (e.g., mouth movements and positions) – are present well before the production of the location in which the hand changes direction and starts to articulate the movement typically associated with the sign, and also well after the location in which the articulation of this movement has come to an end. These facts have been treated as non-linguistic preparatory articulatory activity and subsumed, simply on the basis of tradition, under the notion of transitions, even in the most recent phonetic accounts (e.g., Johnson and Liddell 2011). However, I have suggested in this section that, if we adopt the basic methodological rationale we use to segment words from speech – that the beginning of any phonetic unit is determined by the first occurrence of any articulatory feature that is typically associated with

the unit and, similarly, that the end of any phonetic unit is determined by the last occurrence of the last of such features – then we have to acknowledge that signs are actually longer units than has so far been recognized in mainstream research.

3.2 The perspective of sign recognition

Sign recognition – being able to tell the form and meaning of a sign – is a complex process obviously depending on multiple connections between different levels of grammar and physiology (for a recent overview, see ten Holt et al. 2009). Consequently, even if all the details of this process were known, it would be impossible to go through them in the limited amount of space available here. Fortunately, this is not necessary, since the present section simply aims to discuss two questions: what is the minimum phonetic requirement in terms of sign production for any sign to be successfully recognized, and how far is this consistent with the current mainstream conception of the sign and its phonetic domain? Let us begin, again, by briefly outlining the context when the issue is the recognition of spoken words.

In a simple setting involving a speaker and an addressee (and this setting will be assumed in the remainder of the present section), what is the minimum prerequisite for the addressee to recognize an isolated word produced by the speaker? The almost ridiculously obvious answer is that the speaker must have begun the production of the physical form of the word. That this is indeed the case has been shown in numerous studies at least since Grosjean (1980), whose study indicated that the successful recognition of isolated spoken words required the articulation of at least 30 per cent, and up to 80 per cent, of the form of the words. Without the necessary articulatory information, the recognition of a word can only be explained as guesswork. Guessing, however, does not fit into the current idea of how languages are processed.

Obviously, context affects the recognition rate positively. For example, certain words are more likely to occur together than others, and knowing this generally decreases the time the addressee needs to recognize any given word. However, even in context the production of at least some part of the form of the word needs to be available for the recognizer – if this is not the case, the word is just guessed.

Have studies on sign recognition yielded corresponding results? Certainly it is not always so. Indeed, many studies clearly suggest that signs – produced both in isolation and in context – are recognized already *before* their presently assumed phonetic form has begun to be articulated. One of the earliest of these

studies was the gating experiment conducted again by Grosjean (1981). By artificially modifying the starting place of isolated signs on a video (i.e., gradually allowing viewers to see more and more of the sign) Grosjean found that the recognition of isolated signs required the production of approximately 47 per cent of the form of the sign. The percentage is relatively high, but it must not be taken without reference to Grosjean's method of identifying the beginnings and endings of signs. This reveals that he identified the beginning of the sign with the moment the hands first appeared on the screen, and the end with the moment the hands first started to retract to their resting position. Given this information, the recognition point of signs actually in many cases already occurred during the phase that, in current mainstream terminology, corresponded to the sign initial transitional movement (e.g., ten Holt et al. 2009).

The follow-up study by Clark and Grosjean (1982) with more contextual information indicated that the recognition of a sign in a phrase or sentence might take place even earlier than Grosjean (1981) had found. Clark and Grosjean's study showed that, when the stimuli are excised from continuous signing, only some 37 per cent of the form of the sign needs to be produced for successful recognition. In the study, the beginning of the stimulus sign was associated by a native signer with the moment the previous sign was judged to end. In practice, this too meant that the recognition point of the sign occurred in a large number of cases – again, in terms of mainstream terminology – during the initial transitional movement, not during what is generally taken to be the sign itself.

Subsequent studies in the field have produced similar results (e.g., Emmorey and Corina 1990; Arendsen et al. 2007; ten Holt et al. 2009). In addition to these, a recent study by ten Holt et al. (2009) on isolated signs has shown that the transitional part after the generally accepted *end* of the sign (i.e., the phase during which the hands move back to the resting position) may also be enough for its recognition. In their study, ten Holt and colleagues adopted the gesture-based preparation-stroke-retraction analysis for signs (Kita et al. 1998; Arendsen et al. 2007) and found that stroke phases – the movement phases most strongly associated with the present notion of the sign – and the preparation phases preceding them are alone enough to enable the recognition of a sign. More importantly, they also found that the retraction phase following the stroke was alone sufficient to enable recognition of the sign. There was, though, variation in this particular recognition rate: for example, for the one-handed sign SCHEP 'shovel', articulated in the neutral space in front of the signer, the recognition rate on the basis of the retraction phase alone was near zero, whereas for the one-handed sign TELEFOON 'telephone', articulated with a fairly marked handshape at a fairly marked location on the facial area, the rate was almost 100 percent. Overall, however,

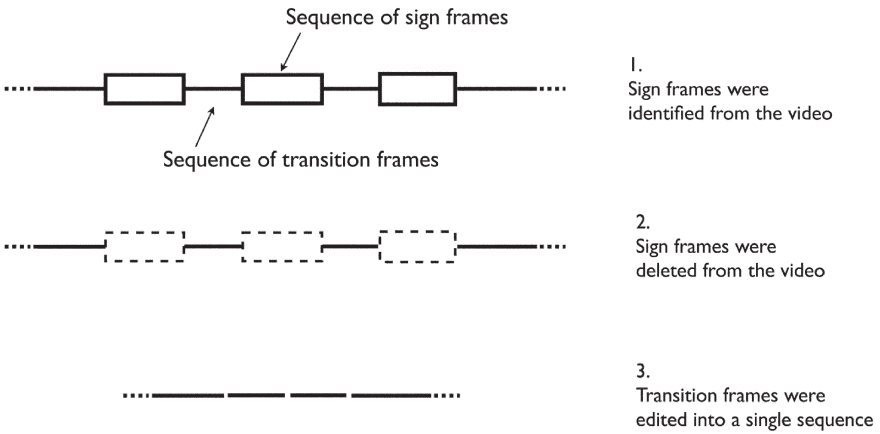

Fig. 7: A schema showing how the video clips without signs were made in Jantunen (2010b).

recognition results based only on the retraction phase were much better than mere chance, with an average of 60 per cent recognition rate for all signs and subjects in their experiment.

Supporting evidence for the study by ten Holt and colleagues (2009) has been provided by Jantunen (2010b). In order to answer the question of whether continuous signing can be understood on the basis of transitions only, he devised a video test in which native Finnish Sign Language signers were shown five short video clips containing superficially continuous-like signing. However, the clips were edited to contain only phases identifiable as inter-sign transitions. This process is described schematically in Figure 7.

The original versions of the five video clips were *Suvi*'s signed examples 500/2, 660/2, 800/2, 860/2, and 1120/2. With the manipulated clips, the task of the signers was to tell whether a clip was understandable or not, and if judged understandable, the signer was asked to repeat the signs in the clip. The main result of the experiment is shown in a simplified format in Table 1.

Overall, the results showed that despite the exclusion of actual "signs", the transition path to and from the location of intended signs was sufficient to allow native signers to construct the intended/target signs; the clips were surprisingly well understood and reproduced on the whole, although some individual variation was found in the performance. The main factor affecting the comprehension/production was found to be the signing speed; a relatively fast signing speed resulting in very short transitions (only one frame) was found to totally prevent the processing in clip 800/2. There were also problems with the comprehension/reproduction of aspect-related semantic details; this was caused by the fact that

Table 1: The main result of Jantunen's (2010b) comprehension/production test. *X* marks that the clip without traditional signs was understood and reproduced correctly.

	Clip 500/2	Clip 660/2	Clip 800/2	Clip 860/2	Clip 1120/2
Signer a	x				x
Signer b	x				x
Signer c	x	x		x	x
Signer d	x				x
Signer e	x	x		x	
Signer f	x	x		x	x
Signer g	x	x		x	x
Signer h	x	x		x	x

these details are typically conveyed by sign-internal local movements which were now missing from the clips.

As far as I am aware, there has been only one study within the recognition paradigm that has directly addressed the question of the location of the borders of signs. This is the early study by Green (1984) that investigated where the intuitions of native American Sign Language signers associate sign borders. The conclusion of the study offered by Green is completely in line with the results of other recognition studies:

> Our results demonstrated that deaf observers do agree on the location of sign boundaries; furthermore, that deaf observers judge signs to begin at the start of the transition movement into the sign [...]. This result suggests the possibility that a sign begins almost immediately after the sign preceding it has ended. It may even be the case that the boundaries of signs can overlap in time, as does the acoustic information for words in speech. (Green 1984: 88.)

The results provided by Green (1984) and, later, other people working within the field of sign recognition, seem to be in direct conflict with claims such as the one made by Johnson and Liddell (2011), that native signers associate the beginnings of signs with the locations in which the movement of the hand appears to change direction, after first having completed the transition to this location. In fact, in the light of the studies presented in this section, such claims begin to appear as mere *post hoc* attempts to justify the use of an empirically invalid, phonologically biased conception of the linear sign and its beginnings and ends. However, regardless of whether this conclusion is accepted or not, this section – together with the previous one – has shown that the question of the empirical borders of the linear sign is not as straightforward as, for example, Johnson and Liddell have proposed. Moreover, this section has suggested that signs may in fact

be longer units than has generally been understood because, according to current mainstream terminology – and on the assumption that some part of the form of the sign needs to be produced before the sign can be recognized – their recognition point is often located within what is now seen as the transitional movement that precedes the sign, and because signs can also be recognized on the basis of parts of their subsequent transitions only.

4 What would the longer sign be like and what consequences might this revision have for sign language research?

The previous section (3) presented phonetic evidence that suggested that the linear sign may be a longer unit than the present mainstream view acknowledges. This section takes up this proposal and discusses, first, what this longer sign would be like and, second, what consequences this "revision" of the sign might have for sign language research.

To begin with, the proposed longer sign would be categorically different from the shorter sign of mainstream tradition. The difference is illustrated schematically in Figure 8 with the help of some structurally different isolated signs from Finnish Sign Language.

For the sake of discussion, let us look more closely at one of these signs, namely the sign BLACK. According to the longer view of the sign BLACK, parts of the sign stream corresponding to traditional transitional movements preceding and following the traditional sign are now included as parts of the sign. These phases are not directly shown in Figure 8, but during the initial phase (before the traditional short sign) the hand is raised toward the location that is depicted in Figure 8 (the traditional beginning location of the sign); the hand orientation and configuration as well as the non-manual element of the sign are identifiable during this phase (in BLACK, the mouth moves in a similar way to the silent pronunciation of the beginning of the Finnish word *musta*, resembling the production of the sound sequence [muss]). During the end phase (after the traditional ending location of the sign), the hand starts to move downwards from the forehead back to the neutral resting position, the orientation and the configuration of the hand as well as the non-manual element gradually neutralizing.

In the longer revised sign, the units that are traditionally termed transitions are actually split in two. In the production of isolated signs such as BLACK in Figure 5, the first transition is re-analyzed as composed of the initial release of the hand plus the phase during which the articulatory features become those

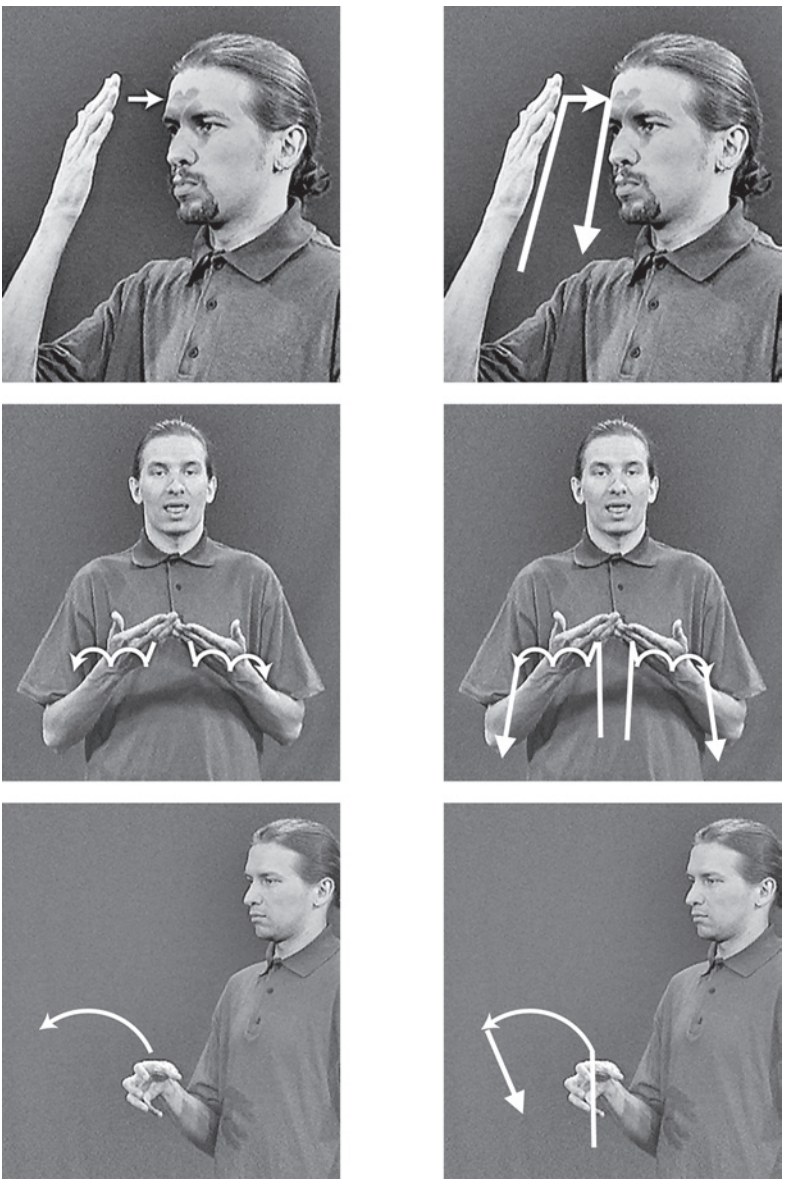

Fig. 8: Traditional short (left) and revised long forms (right) of Finnish Sign Language signs BLACK (one-handed monomorphemic sign; top row), TO-PRESENT (two-handed monomorphemic sign with several sign-internal movement sequences; middle), and TO-GIVE-A-BALL-SHAPED-OBJECT (one-handed semantically complex directional sign; bottom row). Images from the Basic Dictionary of Finnish Sign Language (Malm 1998).

associated with the sign, while the final transition is split into the phase during which the articulatory features become neutralized and the final settling down of the hand. In continuous signing, the very first and very final phases (the initial release and the final settling down) are typically not present. The consequence of this is that continuous signing does not actually contain linguistically redundant activity at all, contrary to the current mainstream view with its notion of transitional movement (Jantunen 2013). Instead, signing proceeds uninterruptedly from sign to sign, without intervening transitions, as does speech, and as already suggested by Green (1984).

To say that, in the longer view of the sign, signing proceeds uninterruptedly from sign to sign is not to say that segmenting signs from each other is easy. Coarticulation causes the borders of signs to overlap, and the borders of signs are always fuzzy. However, although the presence of two manual articulators may cause two signs to be produced simultaneously in a very overt way in sign languages, fuzzy borders as such must not be interpreted as being a property in any way specific to sign language: the issue is exactly the same with words and their boundaries in speech, and, from the perspective of sign language research, would merely require methodological rewiring in order to be taken into account.

In determining the borders of the longer sign, acoustic phonetics carried out with the help of motion capture (mocap) technology may be of great value. When the motion of signs in a signed signal is investigated with the help of mocap, it is possible to see differences in the sign stream that otherwise would pass unnoticed. One such difference concerns the acceleration characteristics of the hand: on the basis of mocap data from continuous Finnish Sign Language signing (Jantunen et al. 2012), it appears that the beginning and ending moments of longer signs correspond remarkably accurately to the moments when the acceleration of the hand is at its highest (acceleration peaks) or at its lowest (deceleration peaks) (for more, see Jantunen 2011, 2012). This is demonstrated with an acceleration curve for the Finnish Sign Language sign 'the house is at place *x*' (see Figure 6 in Section 3.1) in Figure 9. In the figure, the moment when the orientation of the hand first begins to resemble the orientation of the particular sign (frame 3 in Figure 6) is associated with the first full peak (a deceleration peak) in Figure 9 (this moment is marked with a dotted vertical line). Similarly, the moment when the final articulatory feature of the sign has disappeared is equated with the very final (acceleration) peak, or a mound (also marked with a dotted line). The domain of the traditional short sign is identified in Figure 9 with two straight lines in the middle of the curve; the domain includes an acceleration peak and then a deceleration peak, (i.e., the motion of the hand first speeds up and then slows down).

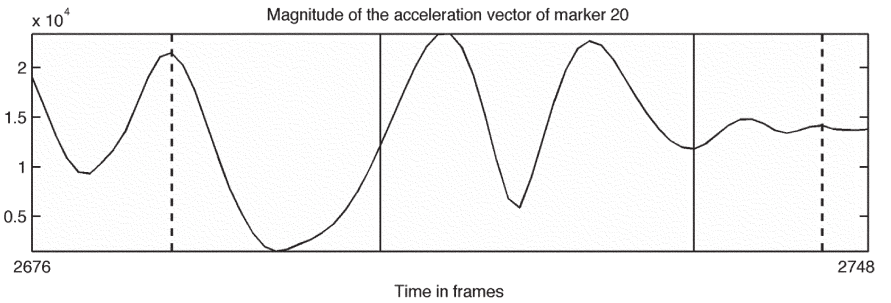

Fig. 9: A motion capture graph describing the magnitude of the acceleration vector of the tip of the dominant hand index finger (y-axis) as a function of time (x-axis). The measurement has been done at the speed of 120 Hz. The domain between the (outer) dotted lines represents the longer form of the contextually produced sign 'the house is at place *x*' (see Figure 6); the smaller domain between the straight lines represents the traditional short version of the sign. Data from Jantunen (2011).

Should the sign be revised into a longer unit, there would be immediate consequences, for example, for sign language phonology and corpus work. Some questions would require answering, including how the revised sign – and especially the movement component – should be represented phonologically, and how these new signs should be identified in the creation of video-based sign language corpora. These questions need to be studied in more detail in the future. However, some initial remarks may already be made here.

Concerning phonology, the fact that the phonetic borders of the sign lie outside the presently acknowledged domain of the sign probably requires some sort of modification of many of the present phonological models. The pressure to do this applies particularly to models that take a stand on the temporal dimension of the sign in the form of describing signs' sequential structure. As was described in Section 2.1, this means practically all the mainstream phonological models (e.g., Brentari 1998; Liddell and Johnson 1989; Perlmutter 1992; Sandler 1989). However, the exact nature of this revision is not yet fully known and should be further investigated in the future. It may be, for example, that the high level of abstraction in the current sequentially oriented mainstream models is also enough to capture the essence of the longer sign, which would mean that the necessary revision of the models would be only minimal. Moreover, and somewhat ironically, in Stokoe's (1960) traditional model, which does not discuss the sequential nature of the sign's structure at all, there might even be no reason to change anything – at least in terms of describing the emic components of the longer sign. Furthermore, there are even "sequential" models in

which movements are not treated as phonological units but as automatic consequences resulting from the fact that, both in isolated signs and in continuous signing, the hand(s) simply "acquire" phonologically specified locations (van der Hulst 1993). In such models, the need for revision would perhaps also be only minimal.

Interestingly, the view of the proposed longer sign corresponds fairly closely to what Kita and colleagues (1998) suggested was the phasal structure of gestures. In their framework, each gesture is typically analyzed as being composed of at least three main components: preparation, stroke, and recovery. *Preparation* is the phase during which the hand moves in a controlled manner toward the starting location of the stroke and during which the orientation and the configuration of the hand come to resemble those typically associated with the gesture. The *stroke* is the main part of the expressive phase of the gesture, and it also corresponds to the traditional short sign of the mainstream view. The final movement phase of the gesture is the *recovery*, during which the shape and orientation of the hand start to relax and the hand starts to retract towards a neutral resting position. In the recovery, the orientation and the configuration of the hand are still identifiable. Kita and colleagues devised their model primarily to segment isolated gestures. However, the model has already been successfully adapted to the study of isolated signs (e.g., Arendesen et al. 2007; Jantunen 2011) and Jantunen (2011) has also used it to segment continuous signing into signs. In the future, the possibility of analyzing signs phonologically according to the model should be investigated further.

The longer view of the sign also has consequences for corpus work on sign languages. Traditionally, the annotation of signs in corpora has assumed that there are always transitional sequences in between signs *proper*, as shown in Figure 4. In the revised, longer view, however, there are no real transitions between signs. This means that there should be no gaps between annotation cells either. What this type of annotation looks like in comparison to traditional annotation with clear transitional phases is shown in Figure 10.

That the role of transitions in corpus annotations should not be exaggerated is a point that has already been partly addressed in some annotation guides (e.g., Johnston 2009). However, as has been shown in this paper, most convention guidelines for annotation still go along with the traditional view of the length of the sign.

It is very likely that, should the sign be considered differently, and along the lines suggested here, there would be consequences too in the more applied domains of sign language research. One example of that would be in the field of sign language teaching which, at least in Finland, has been obsessed with the idea of students learning individual (short) signs, not complete utterances. If it was

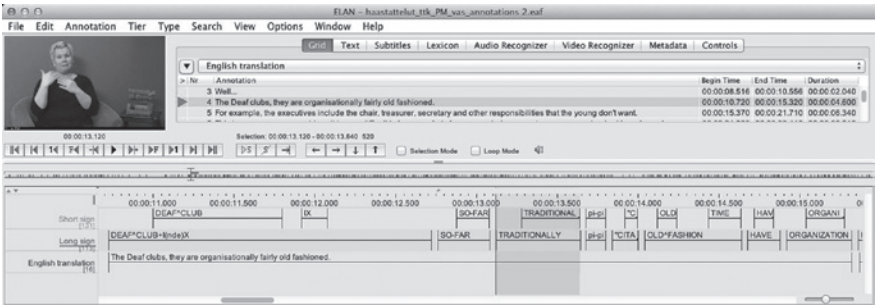

Fig. 10: A screenshot from ELAN (The Max Planck Institute of Psycholinguistics, Nijmegen, The Netherlands; http://tla.mpi.nl/tools/tla-tools/elan/) showing sign annotations that follow both the shorter (upper tier) and the suggested longer view (middle tier) of the sign.

accepted that signs are longer units that melt together in continuous signing, then we would perhaps be able to develop pedagogical methods that would result in many second language learners of sign languages achieving better fluency. However, for the time being, as long as signs are considered to be relative short units, this remains merely a hypothesis waiting to be tested.

5 Conclusion

This paper has discussed the length of the linear sign. It has been shown that the present concept of such a unit derives from early phonological work and it has been suggested here that this concept may in fact be misguided in that it is not based on prevailing phonetic evidence from, for example, the study of the articulation and recognition of signs. Consequently, it has here been proposed that the sign may be a longer unit than is currently recognized in mainstream research. Finally, the nature of this revision has been outlined. In future research, the consequences of this revision in various fields, including for example sign language phonology and corpus work, need to be further investigated.

Acknowledgments: The author wishes to collectively thank the numerous people who have contributed to the completion of the present study and paper. A special thanks is due to Tuija Wainio and to the three anonymous reviewers of the paper for their insightful and helpful comments. The study has been supported financially by the Academy of Finland under grants 134433 and 269089.

References

Arendsen, Jeroen, Andrea van Doorn & Huib de Ridder. 2007. When and how well do people see the onset of gestures? *Gesture* 7(3). 305–342.

Battison, Robin. 1978. *Lexical borrowing in American Sign Language*. Silver Spring, MD.: Linstok Press.

Brennan, Mary. 1980. *Word formation in British Sign Language*. Stockholm: University of Stockholm dissertation.

Brentari, Diane. 1998. *A prosodic model of sign language phonology*. Cambridge, MA: MIT Press.

Clark, Lorene E. & François Grosjean. 1982. Sign recognition processes in American Sign Language: The effect of context. *Language and Speech* 25(4). 325–340.

Crasborn, Onno. 2001. *Phonetic implementation of phonological categories in Sign Language of the Netherlands*. Utrecht: LOT.

Crasborn, Onno. 2012. Phonetics. In Roland Pfau, Markus Steinbach & Bencie Woll (eds.), *Sign language: An international handbook* (Handbooks in linguistics and communication science 37), 4–20. Berlin & Boston: De Gruyter Mouton.

Crasborn, Onno & Inge Zwitserlood. 2008. Annotation of the video data in the *Corpus NGT*. Department of Linguistics & Centre for Language Studies, Radboud University Nijmegen, The Netherlands. http://hdl.handle.net/1839/00-0000-0000-000A-3F63-4 (accessed 15 January 2013).

Emmorey, Karen & David Corina. 1990. Lexical recognition in sign language: Effects of phonetic structure and morphology. *Perceptual and Motor Skills* 71. 1227–1252.

Green, Kerry. 1984. Sign boundaries in American Sign Language. *Sign Language Studies* 42. 65–91.

Grosjean, François. 1980. Spoken word recognition processes and the Gating Paradigm. *Perception & Psychophysics* 28(4). 267–283.

Grosjean, François. 1981. Sign and word recognition: A first comparison. *Sign Language Studies* 32. 98–125.

Hardcastle, William J. & John Laver (eds.) 1997. (eds.) *The handbook of phonetic sciences* (Blackwell handbooks in linguistics). Oxford: Blackwell Publishing.

ten Holt, Gineke, Andrea van Doorn, Huib de Ridder, Marcel Reinders & Emile Hendriks. 2009. Which fragments of a sign enable its recognition? *Sign Language Studies* 9(2). 211–239.

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

Jantunen, Tommi. 2010a. A comparison of two linguistic sign identification methods. In Philippe Dreuw, Eleni Efthimiou, Thomas Hanke, Trevor Johnston, Gregorio Martínez Ruiz & Adam Schembri (eds.), *Proceedings of the 4th workshop on the representation and processing of sign languages: Corpora and sign language technologies* (organized as a part of the 7th Language resources and evaluation conference LREC at Valletta, Malta, May 22–23, 2010), 129–132. Paris: ELRA.

Jantunen, Tommi. 2010b. On the role of transitions in signed language (or: What's wrong with the sign?). Paper presented at the 10th conference on the Theoretical Issues in Sign Language Research (TISLR10), Purdue University, West Lafayette, Indiana (USA), September 30–October 2.

Jantunen, Tommi. 2011. Foneettisen viittoman ala [The domain of phonetic sign]. *Puhe ja kieli* 31(2). 51–65.

Jantunen, Tommi. 2012. Acceleration peaks and sonority in Finnish Sign Language syllables. In Steve Parker (ed.), *The sonority controversy* (Phonetics and phonology 18), 347–381. Berlin & Boston: De Gruyter Mouton.

Jantunen, Tommi. 2013. Signs and transitions: Do they differ phonetically and does it matter? *Sign Language Studies* 13(2). 211–237.

Jantunen, Tommi, Birgitta Burger, Danny De Weerdt & Tuija Wainio. 2012. Experiences collecting motion capture data on continuous signing. In Onno Crasborn, Eleni Efthimiou, Evita Fotinea, Thomas Hanke, Jette Kristoffersen & Johanna Mesch (eds.), *Proceedings of the 5th workshop on the representation and processing of sign languages: Interactions between corpus and lexicon* (organized as a part of LREC 2012 at Istanbul, Turkey, 27 May, 2012), 75–82. Paris: ELRA.

Johnson, Robert E. & Scott K. Liddell. 2011. A segmental framework for representing signs phonetically. *Sign Language Studies* 11(3). 408–463.

Johnston, Trevor. 2009. Guidelines for annotation of the video data in the Auslan Corpus. Department of Linguistics, Macquarie University, Sydney, Australia. http://media.auslan.org.au/media/upload/attachments/Annotation_Guidelines_Auslan_CorpusT5.pdf (accessed 15 January 2013).

Johnston, Trevor & Adam Schembri. 2007. *Australian Sign Language: An introduction to sign language linguistics*. Cambridge: Cambridge University Press.

Kita, Sotaro, Ingeborg van Gijn & Harry van der Hulst. 1998. Movement phases in signs and co-speech gestures and their transcription by human coders. In Ipke Wachsmuth & Martin Froelich (eds.), *Gesture and sign language in human-computer interaction: Proceedings of international gesture workshop*, 23–35. Berlin: Springer.

Liddell, Scott K. 1984. THINK and BELIEVE: Sequentiality in American Sign Language. *Language* 60(2). 372–399.

Liddell, Scott K. & Robert E. Johnson. 1989. American Sign Language: The phonological base. *Sign Language Studies* 64. 195–277.

Malm, Anja (ed.). 1998. *Suomalaisen viittomakielen perussanakirja* [The basic dictionary of Finnish Sign Language]. Helsinki: KL Support Oy.

Mauk, Claude. 2003. *Undershoot in two modalities: Evidence from fast speech and fast signing*. Austin, TX: University of Texas dissertation.

Ohala, John. 1997. The relation between phonetics and phonology. In William J. Hardcastle & John Laver (eds.), *The handbook of phonetic sciences* (Blackwell handbooks in linguistics), 674–694. Oxford: Blackwell Publishing.

Ojala, Stina. 2011. *Towards an integrative information society: Studies on individuality in speech and sign*. Turku: University of Turku dissertation.

Perlmutter, David. 1992. Sonority and syllable structure in American Sign Language. *Linguistic Inquiry* 23. 407–442.

Rainò, Päivi. 2002. Mouthings and mouth gestures in Finnish Sign Language. In Penny Boyes Braem & Rachel Sutton-Spence (eds.), *The hands are the head of the mouth: The mouth as articulator in sign languages*, 41–50. Hamburg: SIGNUM.

Rissanen, Terhi. 1985. *Viittomakielen perusrakenne* [The basic structure of sign language] (Publications of the Department of Linguistics at University of Helsinki 12). Helsinki: University of Helsinki.

Sandler, Wendy. 1989. *Phonological representation of the sign: Linearity and nonlinearity in ASL*. Dordrecht: Foris.

Sandler, Wendy. 1993. A sonority cycle in American Sign Language. *Phonology* 10(2). 243–279.

Stokoe, William C. 1960. *Sign language structure. An outline of the visual communication systems of the American deaf* (Studies in linguistics occasional papers 8). Buffalo, NY: University of Buffalo.

Sutton-Spence, Rachel & Bencie Woll. 1999. *The linguistics of British Sign Language: An introduction*. Cambridge: Cambridge University Press.

Suvi = The online dictionary of Finnish Sign Language. The Finnish Association of the Deaf. http://suvi.viittomat.net (accessed 15 January 2013).

Wallin, Lars, Johanna Mesch & Anna-Lena Nilsson. 2010. *Transcription guidelines for Swedish Sign Language discourse* (Version 1). Department of Linguistics, University of Stockholm, Sweden.

Viitaniemi, Ville, Tommi Jantunen, Leena Savolainen, Matti Karppa & Jorma Laaksonen. 2014. S-pot – a benchmark in spotting signs within continuous signing. In Nicoletta Calzolari, Khalid Choukri, Thierry Declerck, Hrafn Loftsson, Bente Maegaard, Joseph Mariani, Asuncion Moreno, Jan Odijk & Stelios Piperidis (eds.), Proceedings of the Ninth International Conference on Language Resources and Evaluation (LREC'14) (organized at Reykjavik, Iceland, 28–30 May, 2014), 1892–1897. Paris: ELRA.

Zeshan, Ulrike. 2002. Towards a notion of 'word' in sign languages. In R. M. W. Dixon & Alexandra Y. Aikhenvald (eds.), *Word: A cross-linguistic typology*, 153–179. Cambridge: Cambridge University Press.