A corpus-based approach to manual simultaneity

Anna Sáfár and Onno Crasborn

1. Introduction

The most conspicuous modality difference between signed and spoken languages is the fact that in signing the two hands are used as primary articulators. The resulting potential for more simultaneous expression of information has intrigued researchers since the early days of sign linguistics, even though little empirical research has been done until recently (see Vermeerbergen, Leeson, and Crasborn 2007 for an overview). Examples of manual simultaneity have been described for several sign languages and at various levels of linguistic structure (prosody, syntax, discourse). While previous studies have described and analyzed a number of structures that exploit manual simultaneity, large-scale corpus-based analyses of the actual use of these phenomena are yet to be conducted.

Studies aiming to provide grammatical description and analysis may rely on prototypical examples of the structures in question, but the corpus linguist deals with language use "in the wild" and has to identify and understand all instances of the phenomenon under investigation. The annotation process is crucial in any corpus-based approach (Aarts 2002) and clear annotation guidelines are indispensable for reliable and comprehensive data selection in the study of a given structure.

The purpose of this paper is to describe a corpus-based approach to the study of manual simultaneity (in particular manual holds) and our first attempt at implementing it for the Sign Language of the Netherlands (*Nederlandse Gebarentaal*, NGT). After discussing previous research on the subject and the new insights we expect from a corpus-based analysis, we present our annotation guidelines and evaluate them based on a pilot study on the *Corpus NGT*.

Bereitgestellt von | De Gruyter / TCS
Angemeldet
Heruntergeladen am | 16.10.19 13:13

1.1. Manual holds in signed languages

Signed languages make use of the two hands as the most important articulatory channel. Unlike in the case of speech, which is articulated by a single vocal apparatus, the hands and arms are symmetric and independently movable. Since the difference between the left and the right hand is not phonologically distinctive at the lexical level in the sign languages studied to date, signs may be articulated by either hand.[1] As a result, the potential for simultaneous articulation is larger in signed than in spoken languages. Another characteristic of the hands as articulators is that they are visible and can be held in the space in front of the signer. So, as opposed to the auditory signals of speech, parts of signs can be maintained for longer periods.

The characteristics described above give rise to a set of phonetically similar but functionally diverse phenomena that we will refer to as manual spreading.[2] They occur when one of the hands is maintained in space after articulating a sign while the other hand continues to sign. Phonetically, they can be described as the persistence of the final articulatory state of a sign over time, but they seem to serve a wide variety of functions at different levels of linguistic structure, including prosody, syntax and discourse.

At the prosodic level, the non-dominant hand may signal the boundaries of prosodic domains. Nespor and Sandler (1999) describe the spreading of the non-dominant hand in Israeli Sign Language, which may occur rightward or leftward of a two-handed sign, but not across certain prosodic phrase boundaries. The authors interpret this as evidence for the existence of a modality universal hierarchy of prosodic constituents, proposed for spoken languages by Nespor and Vogel (1986) and others. On the other hand, this type of non-dominant hand spreading also shows modality-specific properties, since it enables a phonological domain to be directly perceived (Crasborn 2011). Sandler's discussion of these structures as prosodic suggests that at least in some cases, they may not have any morphosyntactic or discourse function, but merely serve to signal prosodic domains (Sandler 2006).

Syntactic functions of the non-dominant hand are mentioned as early as 1975 by Friedman, who described examples from American Sign Language where the non-dominant hand maintains topic or focus, as well as locative, temporal and pronominal reference (Friedman 1975). The first extensive description of manual simultaneity comes from Miller (1994a, 1994b, 2000), who observed over 200 examples in informal conversations between signers of Quebec Sign Language (LSQ). Miller (2000) describes five types of simultaneous constructions:

Bereitgestellt von | De Gruyter / TCS
Angemeldet
Heruntergeladen am | 16.10.19 13:13

1. Locative constructions (usually involving classifiers)
2. Holds of verbs or predicative adjectives with proposition(s) on the other hand
3. Holds of nouns with proposition(s) on the other hand
4. Simultaneous pronominals and determiners
5. "Oppositive/synthetic" constructions, expressing a contrast of two elements

Miller claims that these structures pose questions to syntactic theory at least in two respects. First, the question arises whether linearity is an underlying or surface property of syntax. Miller suggests that the fact that the syntax of speech is linear, could be due to the properties of the articulators, while the underlying structure may be simultaneous. Conversely, signing may be simultaneous on the surface only, with a linear underlying structure (see also Kremers 2012). Second, the fact that sign-internal morphemic elements can persevere and assume a referential role in a sentence should inform discussions on the interaction of morphology and syntax.

Engberg-Pedersen (1994) focuses on manual simultaneity in what she calls polymorphemic verbs (also referred to by other authors as classifier predicates or verbs of motion and location, see Schembri 2003) and comes to similar conclusions as Miller, namely that the study of such constructions will result in a better understanding of the boundary between morphology and syntax. She also raises the question whether and when simultaneous constructions have an obligatory status, or at least when they are preferred to the consecutive organization of information – questions that remain unanswered to date.

By far the best-known treatment of the subject is that of Liddell (2003, Liddell, Vogt-Svendsen, and Bergman 2007), who adopted a cognitive perspective and focused on the discourse functions of manual spreading. Liddell describes signs that are produced by one hand and then maintained while the other hand continues signing and calls them buoys, because "semantically they help guide the discourse by serving as conceptual landmarks" (Liddell 2003: 223). He distinguishes five types of buoys:

1. Fragment buoy: the meaning of a sign is associated with (part of) its final state of production.
2. Pointer buoy: a pointing sign that "points toward an important element in the discourse" (Liddell 2003: 250)
3. List buoy: numeral signs are held and "provide a physical presence to ordered sets" (Liddell, Vogt-Svendsen, and Bergman 2007: 189)

Bereitgestellt von | De Gruyter / TCS
Angemeldet
Heruntergeladen am | 16.10.19 13:13

4. Theme buoy: the vertical index finger of the weak hand "signifies that an important discourse theme is being discussed" (Liddell 2003: 242)
5. Depicting buoy: a buoy that is part of a depicting space

Liddell analyzes fragment and list buoys in terms of blending, a process whereby these buoys become visible instances of conceptual entities in the discourse. The pointer buoy does not undergo blending but directs attention towards entities. A further type, the "point buoy", was described by Vogt-Svendsen and Bergman (2007), who found that pointing signs can also be held in order to "represent a point in time or space" and "used for visualizing temporal and spatial relations between entities" (2007: 217).

According to Liddell, Vogt-Svendsen, and Bergman (2007), most buoys have fixed forms and express fixed meanings. The exception is the fragment buoy where both form and meaning depend on the preceding sign. As a result, distinguishing fragment buoys from phonological perseveration might pose a problem in some cases (Dudis 2000, cited by Liddell, Vogt-Svendsen, and Bergman 2007). Liddell and colleagues suggest that "if the strong hand acts relative to the weak hand in a meaningful way", this is evidence that a buoy was created (Liddell, Vogt-Svendsen, and Bergman 2007: 209).

Liddell, Vogt-Svendsen, and Bergman (2007) compared buoys in three sign languages: American Sign Language (ASL), Swedish Sign Language (SSL) and Norwegian Sign Language (NSL). They found that all types of buoys described by Liddell (2003) appear in these sign languages, showing similar forms and functions. However, in the case of the theme buoy, they only found tentative evidence for its existence in SSL, and in NSL it seems to be used by younger signers only.

As we have mentioned above, corpus-based investigations of manual simultaneity are rare. A study based on a stretch of continous signing by one signer is Nilsson (2007). Nilsson describes the behaviour of the non-dominant hand in a recorded monologue of nearly ten minutes in Swedish Sign Language and finds that "simultaneity is a key concept and there are several factors contributing to the markedly two-handed impression of signed discourse" (2007:164). She describes the activities of the non-dominant hand as a continuum of eight categories, from inactive to active: being in the lap of the signer, held near the chest, mirroring the activity of the dominant hand at the chest, mirroring, doubling, sign fragment, buoy, and finally dominance reversal, where the non-dominant hand becomes active. Nilsson distinguishes sign fragments from buoys. For her, a sign fragment is

> "a special kind of perseveration, where the non-dominant hand has been part of the production of a two-handed sign, and then remains in that position

Bereitgestellt von | De Gruyter / TCS
Angemeldet
Heruntergeladen am | 16.10.19 13:13

> while the dominant hand continues to produce signs. Whereas phonological perseveration is non-meaningful, sign fragments are meaningful and indicate to the listener who/what the topic of the continued discourse is." (Nilsson 2007:168)."

She finds ten instances of sign fragments in her data, all having their origins in symmetrical two-handed signs. In Nilsson's study, buoys are "independently produced by the non-dominant hand" (Nilsson 2007: 170). She discusses pointer, list (static and sequential), point and theme buoys. In total, there are 19 sequences where buoys are used in her data (1151 signs).

To summarize, examples of manual spreading, a type of manual simultaneity, have been described for several sign languages. Attempts have been made to categorize such constructions, mostly based on the properties of the sign that perseveres. Miller's types are based on grammatical categories, while Liddell's classification seems to be based on the handshape or sign that undergoes spreading.

As mentioned above, both Liddell, Vogt-Svendsen, and Bergman (2007) and Nilsson (2007) discuss the issue that some instances of spreading seem to be "meaningful" while others are purely "phonological". This idea is further supported by Sandler's description of prosodic spreading as a purely phonological phenomenon. While intuitively this seems plausible, meaningfulness is hard to operationalize and detailed analysis of every instance is not feasible for large amounts of data, at least at the first stage of data selection and annotation. We need clear, reliable, and straightforward criteria to select data that can be included for further analysis. Below we first describe the insights we expect from the study of corpora, and then how we approached the problem of efficient yet consistent data selection and annotation.

1.2. A corpus-based approach

As we have seen, there is descriptive evidence for the existence and even strong similarity of simultaneous structures involving manual spreading in several sign languages, and some authors have tried to create taxonomies of such constructions. Yet, large-scale corpus-based analyses are still missing. As a result, we know little about the use and pervasiveness of these phenomena in spontaneous conversation (or in other genres, for that matter) in any language community at large. This in fact applies to most domains of our knowledge of signed languages, simply because sign language corpora have not been available until recently (Johnston 2010).

Bereitgestellt von | De Gruyter / TCS
Angemeldet
Heruntergeladen am | 16.10.19 13:13

As Engberg-Pedersen (1994) has pointed out, we do not know when simultaneous constructions are preferred nor whether they are ever obligatory. Intuitively, it seems likely that these bimanual structures are not obligatory; that a linear formulation, without spreading, is always available, even if simultaneity may be preferred. The question then arises whether there are differences between signers within a community, as well as between different sign languages, in the extent of use of simultaneous manual holds of different types. If there are individual differences between signers, we might ask whether these can be explained by extra-linguistic factors such as handedness; for example, signers who show more ambidexterity may use simultaneity more often. Differences in the frequency of use could also be dependent on genres, in which case one hypothesis could be that simultaneity is more frequently used in planned discourse, due to the more complex coordination of the articulators it requires. For specific types of constructions, large sets of examples will allow an in-depth analysis of the shared properties like phonological, syntactic and semantic context, discourse function, and syntactic structure.

Furthermore, a corpus-based approach may offer new insights in terms of the linguistic description of manual spreading. An exhaustive analysis of all cases of manual holds in a sample of signing (as opposed to focusing on unambiguous cases) may shed new light on characteristics they share and where they differ. A large dataset of many instances of such constructions, both more and less prototypical, may help uncover new and underlying properties that can lead to a more functional categorization and thereby a better understanding of such constructions.

2. Manual spreading in the *Corpus NGT*

In a series of ongoing studies, our aim is to describe manual spreading in actual language use based on data from the *Corpus NGT*. Our goal is to investigate spreading phenomena at the prosodic, morphosyntactic and discourse levels, as well as answer questions about the frequency with which such constructions are used, possible individual differences, and the influence of extralinguistic factors, such as handedness, in their use.

While a large corpus is in principle a treasure trove of information, its potential can only be realized after time-consuming preparation of the data, including glossing, translation and adding other types of annotations. It is possible to automate part of this process, as we will describe below. However, manual annotation is in most scenarios impossible to bypass and

Bereitgestellt von | De Gruyter / TCS
Angemeldet
Heruntergeladen am | 16.10.19 13:13

it involves judgements on a case-by-case basis. In order to make the process of annotation efficient, reliable and replicable, it is indispensable to have a clear definition (what are we looking for?) and operationalization (how to recognize it?) of the phenomena under investigation.

In what follows, we describe our approach to identifying and categorizing spreading phenomena in the *Corpus NGT* using a combination of automatic and manual annotation. We present the results of a pilot study aimed at testing the reliability and efficiency of the annotation process.

2.1. The *Corpus NGT*

The *Corpus NGT* (Crasborn, Zwitserlood, and Ros 2008) consists of roughly 72 hours of video material filmed between 2006–2008. Ninety-two deaf signers were recorded in pairs while in free conversation, discussing proposed topics and retelling stories seen on video. The video recordings are stored and annotated in short sessions (2–5 minutes long in general) using ELAN.[4] For a more detailed description of the corpus see Crasborn and Zwitserlood (2008).

At the time of writing, about 15 hours of the corpus has been glossed, yielding nearly 130,000 glosses in total. Manual activity is glossed on separate right and left hand tiers, so that one-handed signs receive one gloss on the tier corresponding to the hand articulating the sign. Two-handed signs receive two glosses, one on each hand tier. The annotations are precisely time-aligned with the lexical movements of each hand. Thus, for cases where the two hands do not move in synchrony (for example one of the hands of a two-handed sign starts or finishes signing before the other hand), the two glosses for the two-handed sign may have different start and end times. If the end state of a sign is maintained, the gloss ends when one of the parameters (handshape, orientation, or location) starts to change, thus the duration of the gloss indicates the duration of the hold. In practice, it is often a change in handshape, or a relaxation of the fingers, that is taken as the end of a sign, especially since location is often constantly changing because many signs have a path movement of some kind. See Figure 1 for an example of an annotation file as visualized in ELAN.

Bereitgestellt von | De Gruyter / TCS
Angemeldet
Heruntergeladen am | 16.10.19 13:13

Figure 1. Annotation file and gloss alignment

2.2. The annotation process

The aim of the annotation process for this study was twofold. First, we wanted to identify all cases of manual spreading in the selected sessions. Second, we wanted to select those cases where the spreading of the hand is likely to have morphosyntactic or discourse functions, in order to facilitate further analysis by different researchers. The selected data will be further annotated in the future to provide a detailed analysis of the examples, but here we only discuss the first annotation phase aimed at the selection of data.

2.2.1. Identifying manual spreading

We defined manual spreading as follows:

> The full final state of one hand of a one-handed or two-handed sign is maintained and overlaps with one or more subsequent signs articulated by the other hand.

Since the *Corpus NGT* is glossed on separate tiers for the right and the left

Bereitgestellt von | De Gruyter / TCS
Angemeldet
Heruntergeladen am | 16.10.19 13:13

hand with independently time-aligned annotations (see above), manual spreading can be identified as a pattern of overlap between the glosses on the two tiers. Each instance where a gloss on one hand's tier overlapped with one or more different glosses on the other hand's tier were identified and annotated automatically. These automatic annotations were verified by a human annotator, as we will describe below.

2.2.2. Identifying salient spreading

As we have mentioned above, our second goal in this annotation phase was to distinguish between instances of spreading that are likely to be purely prosodic and cases which probably have other, morphosyntactic or discourse functions. Although it is impossible to ascertain without further analysis of any given example whether this is indeed the case, we suggest that there are some cues that do imply such functionality. For the purpose of the present study, we decided to use the term "salient" to refer to spreadings accompanied by such cues, in order to avoid expressions like "grammatical" or "meaningful". By saliency we mean that there is evidence that the signer maintains the end state of a sign intentionally, in the sense that there is evidence that articulatory effort is exerted to do so. We operationalized this idea in terms of dominance reversal and phonological location, as follows:

> A spreading is salient when either of the following is present:
> a) hand dominance is reversed (either to produce the hold or in order to maintain it)
> b) the spreading hand serves as the location that is used by one of the following articulators: the dominant hand (e.g. the location of a pointing sign or a spatially modified sign), eye, head or torso.

The first condition is based on the idea that most signers have a preferred hand in signing and will use this hand whenever possible to articulate signs. Thus, most signers have a "default" hand and using the other hand requires effort, and we assume that in most cases when they make this effort it is for a reason. Since the preferred hand may be difficult to identify in the case of some signers (Sáfár, 2012), we favour a more phonetic definition of dominance reversal, whereby "dominant hand" is defined as the hand that moves while signing and a "reversal" is any point at which there is a dominance switch between the two hands.

In relation to spreading, we assume that when a dominance reversal

Bereitgestellt von | De Gruyter / TCS
Angemeldet
Heruntergeladen am | 16.10.19 13:13

precedes or follows the sign where the spreading originates, this implies that the signer makes an effort to maintain the end state of that sign. An example of spreading followed by dominance reversal is shown in Figure 2. The signer has been signing with his right hand, but after articulating the sign, he maintains its end-state and continues signing with the other hand.

R IF COCHLEAR IMPLANT TAKE-OFF-----
L

if [he] takes the cochlear implant off

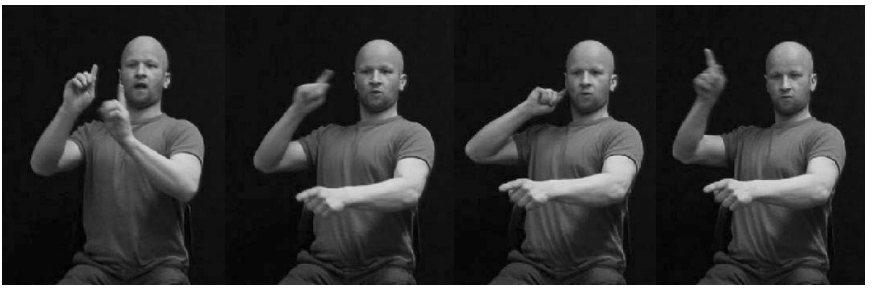
R (TAKE-OFF)----------
L HEAR NOTHING PALM-UP

[he] hears nothing

Figure 2. Spreading followed by dominance reversal (CNGT0529 S026 03:02)

In the example shown in Figure 3 below, the dominance reversal precedes the spreading sign: the previously non-dominant hand (in this case, the left hand) is used to articulate a single sign that is then held in place while dominance reverts to the other hand again. Thus the non-dominant hand becomes active only for the articulation of a single sign which is then held.

Bereitgestellt von | De Gruyter / TCS
Angemeldet
Heruntergeladen am | 16.10.19 13:13

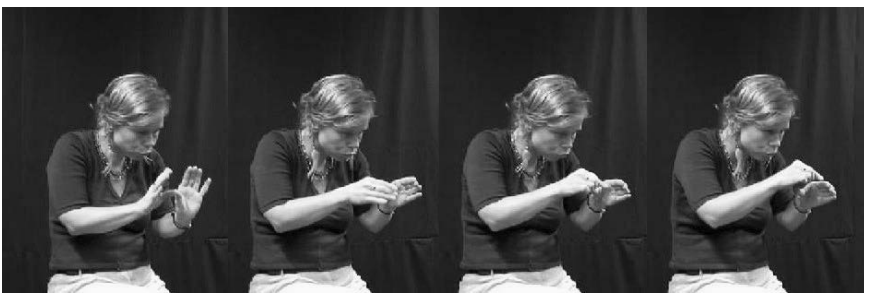


R AS HEARING
L AS PT----------

(they are the same) as the hearing

Figure 3. Spreading preceded by dominance reversal (CNGT0215 S012 00:31)

The second condition in our operationalization of saliency refers to activity from another articulator that uses the spreading hand as its location. In the clearest cases, it is the dominant hand that points to, touches or directs a sign towards the held hand, as shown in Figure 4. However, eyegaze may also be directed to the persevering hand, as well as head or body movements. Such activity is phonetic evidence that there is a morphosyntactic, semantic or discourse relation between the spreading hand and other parts of the utterance. This is similar to Liddell, Vogt-Svendsen, and Bergman's (2007) solution to the problem of distinguishing fragment buoys and phonological perseveration. For them, "if the strong hand acts relative to the weak hand in a meaningful way" (Liddell, Vogt-Svendsen, and Bergman 2007: 209), this is evidence that a buoy was created.



R GROUP PT-arc
L GROUP ----------

Figure 4. Sign directed towards the spreading hand (CNGT0253 S013 08:33)

To summarize, we attempted to formulate criteria to identify spreadings that are likely to have morphosyntactic or discourse functions, based only on information about their phonetic environment. A spreading is selected

Bereitgestellt von | De Gruyter / TCS
Angemeldet
Heruntergeladen am | 16.10.19 13:13

as salient if there is dominance reversal in its immediate environment or if it serves as the location to the activity of another articulator. We hypothesize that these cues suggest grammatical functions of the spreading hand. However, we have to emphasize that the existence of such functions cannot be ruled out in the case of non-salient spreading. Thus, we do not suggest that non-salient spreading has only prosodic functions, but rather that we have no observable evidence for grammatical functions.

2.3. The pilot annotation

In order to evaluate the annotation criteria outlined above, we annotated a subset of the *Corpus NGT*. We wanted to test whether our criteria are unambiguous enough for agreement by independent raters (reliability) as well as whether our distinction between salient and non-salient spreading is in fact useful in distinguishing between different types of spreading.

Data were selected so as to include signers with different characteristics in terms of age, gender, and handedness. Only glossed sessions were selected and both signers were included for each session. In total, 16 sessions involving 22 signers were included in the pilot study (see Appendix 1). All sessions are discussions of issues related to deafness or sign language. The duration of the sessions totals 1 hour and 5 minutes.

Annotations were added on a separate tier by the first author. Each case of automatic annotation (see section 2.2.1.) was evaluated as to whether spreading is indeed present, and if so a saliency judgement was made. For three sessions, a second annotator independently repeated the same process, in order to test the reliability of annotation.

3. Results

3.1. Reliability

Interrater reliability was established for three sessions (CNGT0130, CNGT0253, CNGT0529). For these sessions two annotators independently inspected each instance of spreading found on the basis of existing gloss annotations. They classified each case where a gloss for one of the hands overlapped with one or more non-identical glosses on the other hand as “no spreading”, “spreading” and “salient spreading” (for definitions and criteria

Bereitgestellt von | De Gruyter / TCS
Angemeldet
Heruntergeladen am | 16.10.19 13:13

see 2.2.1. and 2.2.2. above). The first category was necessary because it became clear that the two annotators often disagreed with the annotators who (in the past) had glossed the files. Thus in many cases, they found that they could not see the spreading of a sign over other signs, despite the fact that the glosses showed such relations. In what follows, we first discuss the agreement between the annotators, considering whether spreading was indeed present and if so, whether they were judged to be salient.

In the three sessions (6 signers, 21 minutes, 4215 glosses in total), 175 cases of automatically annotated spreading were examined. In six cases, one of the annotators could not make a decision. Of the remaining 169 cases, annotators were in agreement in 130 cases (77%, Cohen's $\kappa = 0.57$), that is they both judged an instance either as spreading or not spreading. This level of agreement is moderately good, but also shows that there is uncertainty in these judgements.

Concerning the evaluation of the automatic spreading annotations, 106 of the 169 instances (62.7%) were judged as spurious (thus no spreading present) by at least one annotator, and in 67 cases there was agreement between the annotators (40%). Thus, in 40% of the cases of spreading identified automatically based on gloss alignment, both annotators disagreed with the annotator who had added the glosses.

This level of disagreement concerning glossing draws attention to the difficulty of identifying meaningful units in natural signing and uniformly establishing their alignment. When we took a closer look at the cases of spurious spreading, it became clear that in the majority of cases the non-dominant hand was in a rest position that resembled a pointing sign or mirrored the handshape of the active hand.

Regarding the reliability of saliency judgements, that is the decision whether a spreading is salient or not, based on the criteria we outlined above, our results are more promising. In 63 instances, spreading was established by both annotators. Of these instances, 27 (43%) were judged as salient and 32 as non-salient by both annotators (93.65% agreement), while in four cases the annotators did not agree (Cohen's $\kappa = 0.87$). Thus, in the cases where spreading is clearly present, annotators show good agreement in their judgement of saliency. This suggests that our criteria are well-defined, since in the large majority of the cases the two annotators were able to arrive at the same decisions based on these criteria.

The results also give us a first idea as to the usefulness of our criteria. Part of the usefulness of the distinction is that it should result in efficient triage: the criteria should not result in selecting too much or too little of the

Bereitgestellt von | De Gruyter / TCS
Angemeldet
Heruntergeladen am | 16.10.19 13:13

data. For example, our annotation of spreadings as salient or non-salient will not have much significance if 90% of the cases are judged as salient. At this stage, we do not have information about the frequency of spreadings that have prosodic versus morphosyntactic and discourse functions. However, in terms of data selection, the fact that about half of the spreading cases were judged as salient seems to be promising, as this suggests that the criteria we apply result in selecting about half of the cases for further analysis and so the amount of data is considerably reduced.

3.2. Frequency of manual spreading

As described above, we selected 16 sessions (1 hour 5 minutes) for this pilot study. These sessions contain 10,533 glosses from 22 signers. As we have mentioned before, the two hands are glossed separately in the *Corpus NGT*, which means that two-handed signs receive two glosses. It follows that the number of glosses is larger than the number of signs in the dataset. However, a good approximation of the number of signs is the sum of preferred hand glosses (7114). We have listed the number of glosses for each signer in Appendix 2.

In this dataset, we annotated 171 cases of manual spreading, 70 (41%) of which were judged to be salient (see Appendix 3 for the number of spreadings for each participant). This means that on average, for every 100 glosses there are 1.62 spreadings (0.66 salient spreading). However, there are large individual differences between signers in the number of spreadings they use. Figure 5 shows the average number of salient spreadings per 100 glosses for each signer (the raw count for each signer is listed in Appendix 3, but because the number of glosses analyzed per participant varies greatly, it would be deceptive to compare the actual number of spreadings per signer). Four signers did not use salient manual spreading at all, while one signer used salient spreadings very frequently, as many as 7.8 per 100 glosses. The majority of participants showed between 0.5 and 1 salient spreading for every 100 glosses.

Bereitgestellt von | De Gruyter / TCS
Angemeldet
Heruntergeladen am | 16.10.19 13:13

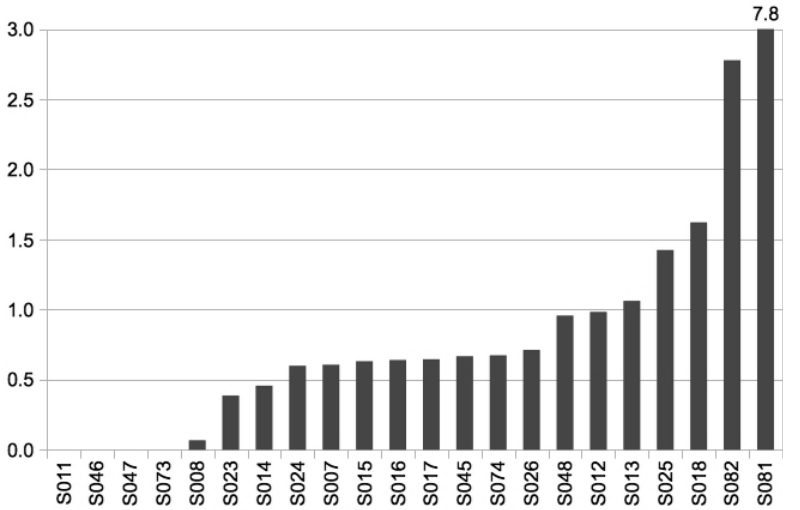


Figure 5. Average number of salient spreading per 100 glosses

These results show that there is great individual variation in the frequency of salient spreadings. However, this could be in part due to the size of the sample: for most signers only a few hundred signs were analyzed. In such short segments, the content of the signing may influence the number of spreadings used, despite the fact that the materials are similar in topic and style.

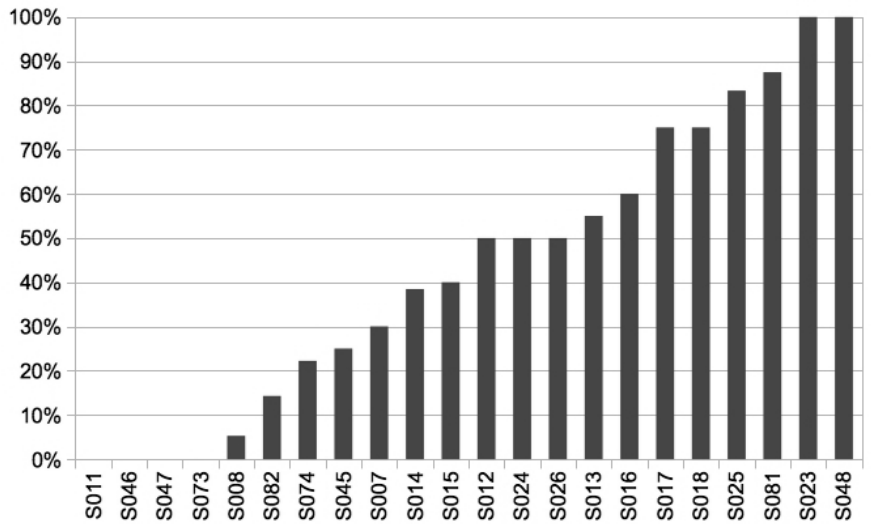


Figure 6. Percentage of salient spreading of all instances of spreading

Bereitgestellt von | De Gruyter / TCS
Angemeldet
Heruntergeladen am | 16.10.19 13:13

Signers also differ widely in the proportion of salient spreading, as Figure 6 shows. For two signers, all instances of spreading are salient, while for other signers only a few cases of manual spreading are salient. Thus there seems to be no direct relationship between the frequency of manual spreading in the signing of a participant and how often these spreadings are judged as salient (that is, likely to have morphosyntactic or discourse functions).

3.3. Usefulness of the criteria

We have pointed out earlier that one of the goals of this annotation phase is to select data to be further analyzed, and thus we expect our criteria to select not too few or too many of the available cases. In the full dataset, 41% of the spreadings were judged to be salient, which is very close to what we saw in the cases annotated for interrater reliability (see section 3.1.). So the "efficient triage" achieved by our criteria applies to the full dataset.

We conclude by taking a closer look at the properties of the constructions selected in the annotation process. We have not conducted syntactic or semantic analyses, so we cannot at this point provide conclusive evidence that our distinction based on what we called saliency indeed distinguished between spreadings that have morphosyntactic and discourse functions and those that are purely prosodic. However, some aspects of the data may provide information about the usefulness of our criteria.

Looking at the cases where spreading was found to be non-salient, we see that all of these spreadings originate in two-handed signs. This may in part be an artefact of the dominance reversal condition. Since one of the criteria for identifying a spreading as salient is dominance reversal in its immediate environment, all one-handed signs that spread will be selected, because maintaining a one-handed sign will require reversing dominance, as the hitherto dominant hand is now held stationary.

However, due to the other condition of saliency, serving as a location for another articulator, some spreadings originating in two-handed signs are also selected as salient. Of the 70 cases of salient spreading found in the data, 12 (17%) spread from a two-handed sign. In this aspect, our approach arrives at results different from Nilsson (2007), who made a distinction between sign fragments (meaningful perseveration from two-handed signs) and buoys (which are always independently produced by one hand). While Nilsson distinguished holds that originate in one-handed versus two-handed signs, our selection criteria results in both types of holds being marked as salient.

Bereitgestellt von | De Gruyter / TCS
Angemeldet
Heruntergeladen am | 16.10.19 13:13

We also looked at how often each condition resulted in selecting a spreading as salient. In 30 cases a spreading was judged as salient because of the presence of dominance reversal, and in 17 cases because the spreading hand served as a location for another articulator. In 23 cases, both conditions were present. This shows that both elements of our definition contribute to selecting salient cases of spreading.

4. Conclusions

In this paper, we described an approach to studying manual simultaneity in a large sign language corpus. Our aim was to develop criteria for reliable and efficient data selection. The annotation process we describe in section 2.2 serves as the first phase in our research project, where we select instances of spreading that will be further analyzed. In this first annotation pass, our aim was to select instances of manual spreading and classify them as salient (likely to have morphosyntactic or discourse functions) or non-salient (purely prosodic).

We tested the proposed method by annotating about one hour of signing from the *Corpus NGT*. The reliability of annotation in terms of inter-rater agreement was established based on a subset of this data. We found that while it is often challenging to identify spreadings (due to the difficulty of interpreting certain handshapes as signs), once a spreading has been identified, the judgement of whether it is salient can be made reliably based on our criteria.

The results of the pilot study suggest that there is individual variation in the use of manual spreading, both salient and non-salient, although these results will have to be replicated in a larger dataset. Based on our criteria, two-handed signs may give rise to both salient and non-salient spreadings, but spreadings from one-handed signs are always judged as salient, due to the fact that saliency is in part defined by the presence of dominance reversal. Further analysis is needed to ascertain whether this is justified. Especially in the case of weakly lateralized signers, who reverse dominance frequently, this may lead to noise in the data.

In the case of two-handed signs, there might be other cues that should be taken into account. For example, we see cases where spreading is interrupted: a spreading hand participates in one or more two-handed signs, but is then returned to its previously held position. This could also suggest syntactic or discourse functions, even in the absence of dominance reversal or activity of other articulators being directed towards the held hand.[5]

Bereitgestellt von | De Gruyter / TCS
Angemeldet
Heruntergeladen am | 16.10.19 13:13

To conclude, we believe our results show that the criteria we proposed for selecting cases of manual spreading that are likely to have morphosyntactic or discourse functions are both reliable and useful in terms of data selection. While further adjustments may be needed based on future analyses, we suggest our approach can be useful in facilitating corpus-based investigations of manual simultaneity. Moreover, our procedure has hinted at a specific problem in glossing sign language texts: where is the boundary between meaningless (yet perceivable) phonetic behaviour and the intentional realisation of a lexical item? Motoric copying behaviour of the non-dominant hand ("echos") and reduced articulations of two-handed signs make it hard to draw the line between these two categories. This appears to be especially problematic for index finger extension, which is physiologically likely to happen as the hand moves to a rest position, yet will not always function as a pointing sign in need of a lexical gloss.

Acknowledgements

This work was made possible by funding from ERC Starting Grant no. 210373 and NWO VIDI grant no. 276-70-012 both awarded to Onno Crasborn. The authors would like to thank Johan Ros, Inge Zwitserlood, Els van der Kooij and Ellen Ormel for discussion of various related issues.

Bereitgestellt von | De Gruyter / TCS
Angemeldet
Heruntergeladen am | 16.10.19 13:13

Appendices

Appendix 1. List of the *Corpus NGT* sessions included in the pilot study.

Session code	Participant 1	Participant 2	Discussion topic	Duration (min:sec)
CNGT0128	S008	S007	deafness	05:59
CNGT0130	S008	S007	deafness	05:31
CNGT0215	S011	S012	deafness	06:23
CNGT0253	S013	S014	deafness	10:14
CNGT0295	S018	S017	deafness	06:37
CNGT0328	S016	S015	deafness	00:42
CNGT0330	S016	S015	deafness	02:36
CNGT0331	S016	S015	deafness	02:07
CNGT0476	S023	S024	deafness	02:57
CNGT0486	S023	S024	sign language	04:23
CNGT0529	S026	S025	deafness	05:21
CNGT1028	S045	S046	sign language	03:46
CNGT1086	S048	S047	sign language	02:31
CNGT1789	S073	S074	deafness	01:11
CNGT1790	S073	S074	deafness	02:28
CNGT2033	S081	S082	deafness	02:16

Bereitgestellt von | De Gruyter / TCS
Angemeldet
Heruntergeladen am | 16.10.19 13:13

Appendix 2. Number of glosses per participant. The total number of preferred hand glosses is 7114.

Participant	Left hand gloss	Right hand gloss	Total
S007	397	593	990
S008	566	917	1483
S011	70	174	244
S012	162	550	712
S013	353	683	1036
S014	374	721	1095
S015	209	108	317
S016	137	332	469
S017	103	362	465
S018	52	133	185
S023	79	180	259
S024	554	281	835
S025	233	118	351
S026	161	401	562
S045	80	220	300
S046	32	139	171
S047	42	155	197
S048	67	142	209
S073	172	78	250
S074	122	175	297
S081	45	45	90
S082	12	24	36
Total	4022	6531	10553

Bereitgestellt von | De Gruyter / TCS
Angemeldet
Heruntergeladen am | 16.10.19 13:13

Appendix 3. The number of salient and non-salient spreadings per participant.

Participant	Total spreading	Non-salient	Salient	Percentage of salient spreading	Total gloss	Spreading per 100 gloss	Salient spreading per 100 gloss
S007	20	14	6	30%	990	2.02	0.61
S008	19	18	1	5%	1483	1.28	0.07
S011	1	1	0	0%	244	0.41	0.00
S012	14	7	7	50%	712	1.97	0.98
S013	20	9	11	55%	1036	1.93	1.06
S014	13	8	5	38%	1095	1.19	0.46
S015	5	3	2	40%	317	1.58	0.63
S016	5	2	3	60%	469	1.07	0.64
S017	4	1	3	75%	465	0.86	0.65
S018	4	1	3	75%	185	2.16	1.62
S023	1	0	1	100%	259	0.39	0.39
S024	10	5	5	50%	835	1.20	0.60
S025	6	1	5	83%	351	1.71	1.42
S026	8	4	4	50%	562	1.42	0.71
S045	8	6	2	25%	300	2.67	0.67
S046	3	3	0	0%	171	1.75	0.00
S047	1	1	0	0%	197	0.51	0.00
S048	2	0	2	100%	209	0.96	0.96
S073	3	3	0	0%	250	1.20	0.00
S074	9	7	2	22%	297	3.03	0.67
S081	8	1	7	88%	90	8.89	7.78
S082	7	6	1	14%	36	19.44	2.78
Total	171	101	70	41%	10553	1.62	0.66

Bereitgestellt von | De Gruyter / TCS
Angemeldet
Heruntergeladen am | 16.10.19 13:13

Notes

1. In NGT (as in many other sign languages), about half of the signs in the lexicon are articulated with one hand and the other half with two hands (see van der Kooij 2002). The dominant hand can be established (as the moving hand) in one-handed and asymmetric two-handed signs. In symmetric two-handed signs, both hands move, but many of these signs may undergo weak drop and be articulated with one hand only (van der Kooij 2001).
2. Manual spreading is not the only type of manual simultaneity. Different lexical items may be articulated on the two hands fully simultaneously (so that both hands move at the same time), but this occurs rarely and only for short durations (except in poetry, see Crasborn 2006). We do not discuss other types of simultaneity, but see Napoli and Sutton-Spence (2011) for an exploration of the extent of simultaneity in signed languages. A sign may be held without simultaneous activity by the other hand, but we do not consider these cases here.
3. We use the terms spreading, hold and perseveration interchangeably in this paper.
4. ELAN (EUDICO Linguistic Annotator) is an annotation tool for the creation of complex annotations on video and audio resources, developed at the Max Planck Institute for Psycholinguistics in Nijmegen, the Netherlands. For more information, see Crasborn and Sloetjes (2008); Crasborn, Hulsbosch, and Sloetjes (2012) and http://tla.mpi.nl/tools/tla-tools/elan/ where the software is available for download.
5. We would like to thank Inge Zwitserlood for this suggestion.

References

Aarts, Jan

2002 Does corpus linguistics exist? Some old and new issues. *Language and Computers* 40 (1): 1–17.

Crasborn, Onno

2006 A linguistic analysis of the use of the two hands in sign language poetry. In *Linguistics in the Netherlands 2006,* Jeroen van de Weijer, and Bettelou Los (eds.), 65–77. Amsterdam: John Benjamins Publishing Company.

2011 The other hand in sign language phonology. In *The Blackwell Companion to Phonology*, Marc van Oostendorp, Colin J. Ewen, Elizabeth V. Hume, and Keren Rice (eds.), 223–240. Oxford: Wiley-Blackwell.

Bereitgestellt von | De Gruyter / TCS
Angemeldet
Heruntergeladen am | 16.10.19 13:13

Crasborn, Onno, and Inge Zwitserlood

2008 The Corpus NGT: an online corpus for professionals and laymen. In: *Construction and exploitation of sign language corpora. Proceedings of the 3rd Workshop on the representation and processing of sign languages,* Onno Crasborn, Thomas Hanke, Ernst D. Thoutenhoofd, Inge Zwitserlood, and Eleni Efthimiou (eds.), 44-49. Paris: ELRA. Available online: http://www.lrec-conf.org/proceedings/lrec2008/workshops/W25_Proceedings.pdf

Crasborn, Onno, Inge Zwitserlood, and Johan Ros

2008 Corpus NGT. An open access digital corpus of movies with annotations of Sign Language of the Netherlands. Nijmegen: Centre for Language Studies, Radboud University Nijmegen. Available online: http://www.ru.nl/corpusngtuk

Crasborn, Onno, and Han Sloetjes

2008 Enhanced ELAN functionality for sign language corpora. In: *Construction and exploitation of sign language corpora. Proceedings of the 3rd Workshop on the representation and processing of sign languages*, Onno Crasborn, Thomas Hanke, Ernst D. Thoutenhoofd, Inge Zwitserlood, and Eleni Efthimiou (eds.), 39-43. Paris: ELRA. Available online: http://www.lrec-conf.org/ proceedings/lrec2008/workshops/W25_Proceedings.pdf

Crasborn, Onno, Micha Hulsbosch, and Han Sloetjes

2012 Linking Corpus NGT annotations to a lexical database using open source tools ELAN and LEXUS. In: *Proceedings of the 5th workshop on the representation and processing of sign languages: interactions between lexicon and corpus, Onno Crasborn*, Eleni Efthimiou, Evita Fotinea, Thomas Hanke, Jette Kristoffersen, and Johanna Mesch (eds.), 19-22. Paris: ELRA.

Engberg-Pedersen, Elisabeth

1994 Some simultaneous constructions in Danish sign language. In *Word-order Issues in Sign Language: Working Papers*, Mary Brennan, and Graham H. Turner (eds.), 73–88. Durham: International Sign Linguistics Association.

Friedman, Lynn A.

1975 Space, time, and person reference in American sign language. *Language* 51(4): 940–961.

Johnston, Trevor

2010 From archive to corpus: transcription and annotation in the creation of signed language corpora. *International Journal of Corpus Linguistics,* 15 (1): 106–131

Bereitgestellt von | De Gruyter / TCS
Angemeldet
Heruntergeladen am | 16.10.19 13:13

van der Kooij, Els

2001 Weak drop in Sign Language of the Netherlands. In *Signed Languages. Discoveries from International Research*, Valerie L. Dively, Melanie Metzger, Sarah Taub, and Anne Marie Baer (eds.), 27–42. Washington, DC: Gallaudet University Press.

2002 Phonological categories in Sign Language of the Netherlands: The role of phonetic implementation and iconicity. PhD Thesis, LOT, Utrecht.

Kremers, Joost

2012 The syntax of simultaneity. *Lingua* 122 (9): 979–1003.

Liddell, Scott

2003 *Grammar, Gesture and Meaning in American Sign Language.* Cambridge: Cambridge University Press.

Liddell, Scott, Marit Vogt-Svendsen, and Brita Bergman

2007 A crosslinguistic comparison of buoys: evidence from American, Norwegian, and Swedish sign language. In *Simultaneity in Signed Languages*, Myriam Vermeerbergen, Lorraine Leeson, and Onno Crasborn (eds.), 187–216. Amsterdam: John Benjamins Publishing Company.

Miller, Christopher

1994a Simultaneous constructions in Quebec sign language. In *Word-order Issues in Sign Language: Working Papers*, Mary Brennan, and Graham H. Turner (eds.), 89–112. Durham: International Sign Linguistics Association.

1994b Simultaneous constructions and complex signs in Quebec sign language. In *Perspectives on Sign Language Structure. Papers from The Fifth International Symposium on Sign Language Research*, Inger Ahlgren, Brita Bergman, and Mary Brennan (eds.), 131–147. Durham: The Deaf Studies Research Unit.

2000 Multi-channel constructions and universal syntax. Unpublished paper presented at TISLR 7, 22–27 July 2000, Amsterdam, the Netherlands.

Napoli, Donna Jo, and Rachel Sutton-Spence

2011 Limitations on simultaneity in sign language. *Language 86* (3): 647–662.

Nespor, Marina, and Wendy Sandler

1999 Prosody in Israeli Sign Language. *Language and Speech* 42: 143–176.

Nespor, Marina, and Irene Vogel

1986 *Prosodic Phonology*. Dordrecht: Foris.

Bereitgestellt von | De Gruyter / TCS
Angemeldet
Heruntergeladen am | 16.10.19 13:13

Nilsson, Anna-Lena

2007 The non-dominant hand in a Swedish Sign Language discourse. In *Simultaneity in Signed Languages*, Myriam Vermeerbergen, Lorraine Leeson, and Onno Crasborn (eds.), 163–185. Amsterdam: John Benjamins Publishing Company.

Sáfár, Anna

2012 Ambidextrous signers in the Corpus NGT. In: *Studies in applied linguistics in honour of Edit H Kontra*, Éva Illés and Tamás Eitler (eds.), 123-132. Budapest: ELTE BTK.

Sandler, Wendy

2006 Phonology, phonetics, and the nondominant hand. In *Papers in Laboratory Phonology: Varieties of Phonological Competence,* Louis Goldstein, Douglas H. Whalen, and Catherine Best (eds.), 185–212. Berlin: Mouton de Gruyter.

Schembri, Adam

2003 Rethinking 'classifiers' in signed languages. In *Perspectives on Classifier Constructions in Sign Languages*, Karen Emmorey (ed.), 3–34. Mahwah, NJ: Lawrence Erlbaum Associates.

Vermeerbergen, Myriam, Lorraine Leeson, and Onno Crasborn

2007 Simultaneity in signed languages: A string of sequentially organised issues. In *Simultaneity in Signed Languages*, Myriam Vermeerbergen, Lorraine Leeson, and Onno Crasborn (eds.), 1–25. Amsterdam: John Benjamins Publishing Company.

Vogt-Svendsen, Marit, and Brita Bergman

2007 Point buoys: the weak hand as a point of reference for time and space. In *Simultaneity in Signed Languages*, Myriam Vermeerbergen, Lorraine Leeson, and Onno Crasborn (eds.), 217–236. Amsterdam: John Benjamins Publishing Company.

Bereitgestellt von | De Gruyter / TCS
Angemeldet
Heruntergeladen am | 16.10.19 13:13

Bereitgestellt von | De Gruyter / TCS
Angemeldet
Heruntergeladen am | 16.10.19 13:13