

ACCEPTABILITY JUDGMENTS IN SIGN LINGUISTICS

VADIM KIMMELMAN (UNIVERSITY OF BERGEN)

1. INTRODUCTION

Sign languages are natural languages existing in the visual modality and used primarily by deaf people (Pfau *et al.* 2012). Over the past 50 years, much research on sign languages from all around the world has demonstrated fundamental commonalities between signed and spoken languages, such as comparable or identical phonological, morphological, syntactic, semantic, and pragmatic processes and structures, leading to the claim that generally, grammatical concepts and models are modality independent, that is, they can be applied to all natural languages, no matter whether they are spoken or signed. However, the research has also discovered important differences known as *modality effects* (Meier 2012; Sandler & Lillo-Martin 2006). Investigating sign languages and uncovering such modality effects are crucial for our understanding of linguistic universals and the human linguistic capacity.

In order to frame the following discussion, I need to introduce some of the crucial differences between signed and spoken languages (see Meier (2012) for an extensive discussion, including other modality effects that I will not address). First, sign languages are highly iconic on many levels. Many individual lexical signs are iconic: the signs resemble their referents (Taub 2012). But, more importantly, many grammatical features, such as marking of number and aspect, verbal agreement, and verbal expression of motion and location, among others, are also iconic (Davidson 2015; Liddell 2003; Wilbur 2008). For instance, the verb ${}_1\text{CL}(\text{ROUND})\text{-GIVE}_a$ in example (1)¹ and Figure 22.1, from Russian Sign Language (RSL), has several iconic components: the handshape reflects the fact that the given object (the ball) is round, and the movement and orientation of the sign depict the fact that the ball is transferred from the signer to a person to the right of the signer; in addition, the fact that the movement occurs only once reflects that the event occurred only once.

- (1) INDEX₁ BALL ${}_1\text{CL}(\text{ROUND})\text{-GIVE}_a$ [RSL]
'I gave him a ball.'



¹ See Appendix A for the glossing conventions.

exists, researchers use various approaches, and sometimes favor the use of a particular method. I argue that there are clear indications that methodological research is necessary, as often contradictory descriptions of the same phenomena can be found in the literature (Section 3). Given modality-related and socio-linguistic peculiarities of sign languages, I propose which questions should be addressed in future methodological studies (Section 4). Finally, I provide some recommendations for using acceptability judgments as a tool in sign linguistics at the moment – until further methodological research delivers new insights (Section 5).²

2. CURRENT PRACTICE IN THE FIELD

Browsing the literature in the field of sign linguistics, it is quite easy to see that acceptability judgments by native signers of various types are commonly used. Specifically, researchers have used:

- phonological judgments (well-formedness of phonologically modified signs (Arendsen *et al.* 2010) or determining where stress falls (Wilbur & Schick 1987));
- morphological judgments (well-formedness of morphologically modified signs (e.g. Krebs, Wilbur & Roehm 2017));
- syntactic judgments (well-formedness of sentences: many examples, see below);
- semantic and pragmatic judgments (e.g. whether a certain utterance can be used to express a certain meaning in a given context (e.g. Davidson 2013)).

Given the scope of this handbook, in the rest of the chapter, I mostly focus on syntactic acceptability judgments, which are incidentally also the most common type.

Despite the fact that acceptability judgments are often used, close to no methodological research discussing them and/or comparing them to other methods exist.³ A rare exception is Chapter 2 in Neidle *et al.* (2000), where the authors discuss the methodology they used to describe the syntax of American Sign Language (mostly syntactic acceptability judgments) and provide insightful instructions on how to use this method. Note, however, that these recommendations are based purely on the authors' experience and observations and not on any systematic quantitative analysis.

It is instructive to investigate how the collection of judgments is commonly handled. To determine the recent common practice in theoretical sign linguistics, I have skimmed through

² This chapter is only devoted to the use of acceptability judgments. Going beyond eliciting judgments, eliciting data by means of visual stimuli is preferred in order to avoid influence from spoken language – but this method has limitations as not all phenomena are easily elicited with pictures, and no negative data can be elicited. Corpus data is often considered ideal/most natural, but many structures are not or scarcely attested in corpora, and again no negative data is directly available. I discuss the use of corpora in comparison to eliciting judgments later in the paper. For general recommendations on methodology in sign linguistics, see e.g. Orfanidou *et al.* (2015).

³ Note that I am not claiming that no methodological research on sign languages exist: see, for instance, van Herreweghe & Vermeerbergen (2012) and Orfanidou *et al.* (2015). However, for some reason, existing methodological studies hardly ever discuss acceptability judgments, focusing instead on other methods of data collection, ethical issues, etc.

all the research papers published in the journal *Sign Language & Linguistics* in the years 2010-2017. It turned out that 9 out of the 28 papers used acceptability judgments, and only 3 of those 9 applied experimental methods and statistics. Thus using judgments is a common method, but the most common approach to collecting judgment data is informal (as opposed to experimental), which goes against the current recommendations for conducting judgment elicitation (Schütze & Sprouse 2014). Given the recent findings that informally collected judgments are highly correlated with judgments collected using experimental methods (Sprouse *et al.* 2013), this might not be a serious shortcoming. However, whether this can also be generalized to using judgments in sign language research is questionable, as will be discussed in the following sections.

Note that this sample of articles from *Sign Language & Linguistics* is very small and hardly representative of the field of theoretical sign linguistics. Upon further inspection of the publications in the field, it appears that more recent research tends to use experimental methods more often. Several recent studies used an experimental setup to elicit judgments and applied statistical analysis to interpret the findings (Börstell 2017; Davidson 2014; Gökğöz 2013; Kimmelman 2018; Loos 2017).

Recently, a group led by Philippe Schlenker has been using a new method of collecting syntactic and semantic judgments, which they call the ‘playback method’ (e.g. Schlenker 2014; Schlenker 2017). In this method, a native signer is asked to sign a paradigm of sentences (minimal pairs or sets of sentences different in the aspect relevant to the research question, including some which are expected to be ill-formed); these sentences are recorded on video. The resulting videos are later shown to the same signer who is asked to rate the sentences on a 7-point scale; multiple judgments on the same sentences are collected on multiple occasions. It seems useful to conduct methodological research to establish validity and reliability of this method, as I discuss in Section 4.

Moving from theoretical sign linguistics to other topics, in psycholinguistic research, acceptability judgments are commonly used, and in an experimental setup accompanied by statistical analysis. Specifically, acquisition of sign languages and comparison of early vs. late learners are topics which have been extensively studied using syntactic judgments (Boudreault & Mayberry 2006; Cormier *et al.* 2012; Henner *et al.* 2016; Novogrodsky *et al.* 2017). Note that in these studies, the judgment task is used as a tool to assess the level of competence in a sign language rather than to gain insight into the structure of the language.

In most studies that use either informal or experimental collection of acceptability judgments, no space is devoted to discussing methodological issues connected to this method or to comparing judgments with other methods such as corpus research (some exceptions are discussed in the next section). One of the questions that has been widely discussed for spoken languages is whether experimental methods are necessary when collecting judgments; there are researchers arguing both for and against reliability of using informal methods (Featherston 2007; Sprouse *et al.* 2013). For sign languages, no such research has been conducted. However, as I show in the next section, there are indications that current practices of collecting judgment

data are not completely reliable and thus should be evaluated by conducting methodological studies.

3. WHY METHODOLOGICAL STUDIES ARE NEEDED

It is trivial that theoretical research must be based on a good empirical fundament. If the primary data on which a theory is based is faulty, it might completely undermine the validity of that theory. However, in the field of sign linguistics, it is not difficult to find examples of elaborate theories, e.g. formal analyses with important general implications, based on contested data. Let's consider two examples.

One important domain in which sign languages are different from spoken languages is the structure of *wh*-questions. While in spoken languages *wh*-words either move to the left periphery of the sentence or stay in situ, in many sign languages *wh*-signs have been shown to either occupy a clause-final position or to occur twice (Cecchetto 2012). Somewhat simplifying, in generative approaches, the spoken language data is explained by postulating that *wh*-movement targets the specifier of the CP layer, that this specifier can only be on the left, and that therefore only leftward *wh*-movement exists (Kayne 1994). If sign languages in fact have the specifier of the CP layer on the right and allow for rightward *wh*-movement, the general theory of *wh*-movement has to be revised, and the peculiar behavior of sign languages must be explained. In addition, the possibility to double the *wh*-sign has to be explained as well. However, the relevant data from at least some sign languages is contested.

For instance, two different analyses of *wh*-questions in American Sign Language (ASL) exist.⁴ Neidle et al. (1997, 1998) argue that the specifier of CP is on the right and that *wh*-signs in ASL indeed move to the right. This implies the necessity to revise the theory of *wh*-movement in general. In contrast, Petronio & Lillo-Martin (1997) argue that the specifier of CP is on the left and that *wh*-signs in ASL move to the left. They claim that when a *wh*-sign occurs in the right periphery of a sentence, it is a double of the original *wh*-sign base-generated in the C head (which is on the right). Given this line of reasoning, the theory of movement developed for spoken languages also applies to sign languages.

Clearly the two analyses make different predictions. Specifically, there are three crucial test cases: an object *wh*-sign in the left periphery of the sentence (3a), a subject *wh*-sign in the right periphery of the sentence (3b), and a phrasal *wh*-constituent in the right periphery of the sentence (3c). Neidle et al. (1997, 1998) predict that (3a) is ungrammatical, while (3b) and (3c) are grammatical; Petronio & Lillo-Martin (1997) make the exact opposite predictions.⁵

⁴ For the sake of space, my recap of the debate on *wh*-questions in ASL is unavoidably an oversimplification. For instance, I do not discuss and do not gloss non-manual marking. Please consult the original articles and references therein for all the details.

⁵ (3c) should be ungrammatical according to Petronio & Lillo-Martin (1997) because the phrasal *wh*-constituent cannot occur in the C head; (3b) is ungrammatical for a subtler reason: while the sign WHO can occur in the C head, the ungrammaticality is caused by the absence of the original WHO in situ or in the left periphery.

Intriguingly, their findings are in agreement with their own predictions, but in disagreement with each other.⁶

- (3) a. WHO JOHN HATE? [ASL]
‘Who does John hate?’
Neidle et al. (1998 p. 822): *; Petronio & Lillo-Martin (1997 p. 50): OK
- b. HATE JOHN WHO?
‘Who hates John?’
Neidle et al. (1998 p. 822): OK; Petronio & Lillo-Martin (1997 p. 36): judgments vary
- c. JOHN BUY YESTERDAY WHICH COMPUTER?
‘Which computer did John buy yesterday?’
Neidle et al. (1998 p. 827): OK; Petronio & Lillo-Martin (1997 p. 38): *

Both groups of researchers used acceptability judgments provided by multiple native signers administered informally. The differences in their findings indicate that the reliability of this method must be questioned. Furthermore, both groups discuss far-reaching theoretical implications of their findings, which makes the necessity for reliable data ever more obvious.

As mentioned above, another common pattern in wh-questions (also attested in ASL) is that the wh-sign is doubled (i.e. occurs twice in a sentence, usually clause-initially and clause-finally). Interestingly, for Italian Sign Language (*Lingua dei Segni Italiana*, LIS), using informal elicitation of judgments, Cecchetto et al. (2009) found that such doubling of wh-signs is not possible. A follow-up corpus study conducted by some of the same researchers, however, uncovered that naturalistic data contains multiple clear examples of this construction, as in (4) (Branchini *et al.* 2013). Branchini et al. (2013) also collected acceptability judgments for such sentences from two native signers, who this time found them to be acceptable.

- (4) WHAT INDEX₁ DO WHAT? [LIS]
‘What shall I do?’ (Branchini *et al.* 2013 p. 168), compare to Cecchetto et al. (2009 p. 285)

In addition to, once again, casting doubts on the informal elicitation of judgments, the LIS case also points to the phenomenon of possible discrepancies between judgments and corpus data. With the recent emergence of annotated corpora of several sign languages, it becomes possible (and necessary) to further investigate the relation between judgments and naturally produced data, as will be discussed in Section 4.

More examples of contradictory descriptions based on (informally elicited) acceptability judgments, leading to contradictory theories, can be found in the literature – compare, for instance, Wilbur (1996) to Hoza et al. (1997) and Caponigro & Davidson (2011) on the status

⁶ The two sources do not always discuss the exact same sentences, but they discuss structurally identical ones. For instance, (3a) is an example from Neidle et al. (1998 p. 822), and Petronio & Lillo-Martin (1997 p. 50) include a structurally identical example WHO YOU LIKE? ‘Who do you like?’ as grammatical.

of pseudo-cleft-like constructions in ASL, or Neidle et al. (2000) to Abner (2013) on possessive constructions in ASL. Most of these examples concern ASL, but this is due to the fact that this sign language is by far the best studied. For many sign languages, only single descriptions of most phenomena exist, and in the absence of methodological research or replication studies, there is no way to determine whether these descriptions are trustworthy.

Since I only discuss a few isolated examples of contradictory judgments reported in the literature, I can by no means draw the conclusion that informal syntactic judgments in sign linguistics are unreliable in general. It might very well be the case that, similar to what has been found for spoken languages (Sprouse *et al.* 2013), once we experimentally test a large sample of informally collected judgments reported in the literature, we will find them to be reliable. What I argue instead is that the contradictory judgments observed in the literature provide enough motivation to conduct such a methodological study for sign languages.

4. FUTURE RESEARCH DIRECTIONS

Given the lack of methodological research, in the remainder of the chapter, instead of summarizing the literature, I will propose possible research topics that should be investigated, and provide some arguments for the need to investigate them. I start from the general methodological questions and then turn to sign language specific issues in need of further study.

4.1 GENERAL METHODOLOGICAL STUDIES

As discussed in the previous section, a general quantitative research on the reliability of acceptability judgments in sign linguistics is necessary. This could be a replication of the study by Sprouse et al. (2013): one can take a sample of reportedly grammatical and ungrammatical sentences from theoretically oriented sign linguistics papers and test their acceptability using a large sample of signers and proper experimental design. This type of research almost necessarily must use data from ASL, as for other sign languages either not enough published research exists, or it would be difficult to find a sufficiently large sample of native signers (see also the discussion in Section 4.3 below).

An alternative approach to testing reliability of acceptability judgments is not to use judgments from the literature, but (1) to collect judgments both experimentally and informally in a single study⁷ or (2) for experimentally collected data, to investigate within and between participant reliability, as for instance in Langsford et al. (2018). The advantage of this approach is that it can be applied to any sign language – even a relatively poorly studied one.

As discussed in the previous section, some researchers have used the ‘playback’ method (Schlenker 2014), where a paradigm of potentially grammatical and ungrammatical sentences is signed by a native signer and later assessed by the same signer. It is clear that the generalizability of the findings gathered by using this method is questionable. Even if the

⁷ This might prove difficult as in collecting judgments informally the researcher has many choices, each of which can potentially influence the outcome. Judgments reported in the literature come from different authors using different methodological decisions, and thus they are a better representation of using informal collection methods in general.

purpose of the researcher is not to generalize beyond the single signer, the method's reliability within participant should be investigated (note that inconsistent judgments of a single sentence at different time points are sometimes reported). On the other hand, the method is attractive as it is relatively simple to use and also transparent: the sentences and the judgments are recorded and available for inspection. A potential extension of this method is to use a small number of signers (e.g. five) instead of just one, who do not only assess their own production (Myers 2009). Therefore, it would surely be beneficial to conduct a methodological study of this method and to compare it to other informal and formal methods of collecting acceptability judgments.

Finally, it would be very instructive to conduct studies comparing judgments and corpus data (see also Chapter 28 in this volume). Research on spoken languages has demonstrated that judgments are not always consistent with actual language production as reflected by corpora (see e.g. Divjak (2017) and references therein).

Recently, relatively large annotated corpora of several sign languages have become available: NGT: <https://www.ru.nl/corpusngtuk/introduction/welcome/> (Crasborn *et al.* 2008), Australian Sign Language: <http://www.auslan.org.au/about/corpus/>, British Sign Language (BSL): <http://bslcorpusproject.org/data/>, LIS (Geraci *et al.* 2011); a large corpus of German Sign Language is presently being created: <https://www.sign-lang.uni-hamburg.de/dgs-korpus/index.php/welcome.html>. Smaller corpora are also available for other sign languages, including RSL: <http://rsl.nstu.ru/> (Burkova 2015).

Importantly, several corpus-based studies report two types of findings: (1) structures which were previously reported to be ungrammatical are found to be present in the corpus, and (2) much more variation is attested in the corpus than would be expected based on published findings. One such study on wh-doubling in LIS was discussed above in Section 3 (Branchini *et al.* 2013). Another clear example comes from NGT. Oomen & Pfau (2017) study negation based on the corpus of NGT and find that, contrary to previous claims in the literature based on judgments (Van Gijn 2004), a manual negative sign is often present in negative sentences, and also that there is more variation in word order and the scope of non-manuals than has been reported. More examples of corpus data contradicting previous findings and demonstrating much variation can be found in recent publications (Couvee & Pfau 2018; de Beuzeville *et al.* 2009; Fenlon *et al.* 2018).

Corpora are generally considered a very reliable data source in spoken language research, as they reflect real language use well. However, due to the extremely time-consuming process of annotating sign language data leading to much smaller corpora than available for many spoken languages, and also due to socio-linguistic factors discussed in Section 4.3 below, the data in sign language corpora might not be as representative. Furthermore, small corpora especially cannot be used to provide negative data. It is therefore clear that a careful study of the differences between data obtained via acceptability judgments and data present in corpora of sign languages, as well as of the causes of these differences, is necessary (see also Kimmelman *et al.* (2018) for some discussion).

4.2 STUDIES OF POSSIBLE MODALITY EFFECTS

As discussed in Section 1, sign languages are to some extent shaped by the visual modality and are thus in some respects different from spoken languages. The two modality effects that might influence acceptability judgements are iconicity and simultaneity.

4.2.1 ICONICITY

One important modality effect is the abundance of iconicity (Taub 2012). Extensive research on the role of iconicity at the lexical level for sign language acquisition and processing exists (Perniss *et al.* 2010). In second language acquisition, lexical iconicity unsurprisingly facilitates learning; and, despite some controversy, there is also some evidence that it plays a role in first language acquisition (Thompson *et al.* 2012). Furthermore, it has been shown that lexical iconicity plays a role in various processing tasks (Perniss *et al.* 2010).

More relevant for the possible effects of iconicity on syntactic/semantic acceptability judgments is existing research on effects of grammatical iconicity on perception. Recall example (1), repeated here as (5), see also Figure 22.1. The predicate ${}_1\text{CL}(\text{ROUND})\text{-GIVE}_a$ in this example is a so-called classifier predicate (Zwitserslood 2012). It has been shown that all phonological components of such signs (handshape, location, and movement) are iconic. For instance, the handshape iconically depicts the round shape of the object BALL in (5).

- (5) INDEX₁ BALL ${}_1\text{CL}(\text{ROUND})\text{-GIVE}_a$ [RSL]
'I gave him a ball.'

Crucially, the iconic components have the potential of being gradually modified. For instance, the handshape and the distance between the hands depicted in Figure 22.1 suggest a certain size of the moved object, and a slightly smaller handshape and distance might refer to a slightly smaller size, and a slightly larger handshape and distance to a slightly larger size, etc. The same is true for location and movement: minute changes can be meaningful. Interestingly, some research in the categorical perception paradigm shows that, in ASL, at least iconic handshapes, despite being potentially gradual, are perceived categorically by native signers (Emmorey & Herzig 2003). Other research on BSL, however, finds evidence of both categorical and gradual perception of such handshapes (Sehyr & Cormier 2015).

The main question that needs investigation is how iconicity influences acceptability judgments. More specifically, it would be useful to study how iconicity violations are judged by native signers. For instance, we can modify the verb in example (5) in two ways, as shown in (6). In (6a), the first location is incorrect: the verbal sign could be interpreted as 'person B is giving a round object to person A', but not as 'I give a round object to person A'. This can be analyzed as a grammatical violation of verbal agreement with the first-person subject, and an iconicity violation: the first location and the movement of the sign do not correctly depict the movement of the ball. In (6b), the handshape is incorrect: it refers to a square object, while the theme object is in fact round. This can be analyzed as a grammatical violation of verbal agreement with the object, and an iconicity violation: the handshape does not resemble the object it refers to.

- (6) a. INDEX₁ BALL_bCL(ROUND)-GIVE_a [RSL]
b. INDEX₁ BALL₁CL(SQUARE)-GIVE_a

Consequently, the examples in (6) involve a combination of a grammatical and an iconic violation. It is also possible to imagine cases of iconicity violations without a grammatical violation. For instance, in (7a) the verb moves not from location *a* but from the neutral space. Grammatically, this is not a mistake, because subject agreement in sign languages is almost always optional (Lillo-Martin & Meier 2011). However, this is less iconic than the alternative in (7b) and thus constitutes an iconicity violation.

- (7) a. INDEX_a BALL CL(ROUND)-GIVE_b [RSL]
b. INDEX_a BALL_aCL(ROUND)-GIVE_b [RSL]
'She/he gave her/him a ball.'

Modifications of iconic components can also be made subtler: instead of the location or handshape clearly not matching the relevant subjects or objects, a small gradual modification of the sign might be applied. Again, this might lead to an iconicity violation without necessarily making the sentence ungrammatical. Note that some research indeed reports semantic and syntactic effects of gradual modifications of iconic components of signs (Kuhn & Aristodemo 2017; Schlenker 2014).

At the moment, it is not clear to me what to expect with respect to judgments on such iconicity violations in comparison to grammatical violations in non-iconic structures. On the one hand, iconic violations might be easier to spot and thus easier to judge as unacceptable. On the other hand, if gradual perception is at play for such constructions, subtle violations of iconicity might not be perceived as problematic as clear violations of grammaticality in non-iconic structures. Finally, it is not clear whether acceptability of iconicity violations can or should be elicited using the same task formulations as grammaticality violations. Given that many grammatical structures in sign languages also have an iconic component, this might call for a special attention to the creation of acceptability judgment experiments. I thus conclude that there is a need for research that specifically addresses acceptability judgments for iconic structures.

4.2.2 *SIMULTANEITY*

Another important modality effect is the amount of simultaneity in sign languages. Components of signs are combined (mostly) simultaneously; morphemes are almost always combined simultaneously; finally, a lot of grammatical information is expressed via non-manual markers occurring in parallel with manual signs (Pfau & Quer 2010). I argue that it is reasonable to expect at least two types of effects of simultaneity on acceptability judgments. First, it might be the case that judgments on longer sentences might be more challenging and consequently less consistent than would be the case for spoken languages. Second, judgments on word/sign order in general might be less reliable.

The motivation for these possible effects comes from psycholinguistic research. Several studies have shown that working memory capacity for signs is smaller than for words in sequential

recall tasks (Boutla *et al.* 2004; Geraci *et al.* 2008; Wilson & Emmorey 1997).⁸ This is most likely explained by the amount of simultaneous information in signs, and by the less important role serial order plays in the visual modality (*ibid.*). If working memory capacity for signs is smaller, it is natural to expect that longer sentences – which are more difficult to process even in spoken language – would be even more challenging in an acceptability judgment task. Moreover, it is reasonable to hypothesize that all sign order judgments might be less stable. In fact, some researchers have proposed that sign order in general plays a lesser role in sign languages (Bouchard & Dubuisson 1995). However, a common acceptability judgment task used in psycholinguistic research is fully based on sign order violations (Boudreault & Mayberry 2006; Novogrodsky *et al.* 2017).

I thus suggest that it would be useful to test reliability and variation in acceptability judgments concerning word order and longer sentences. Judgments on structures concerning seriality should be compared to judgments on other grammatical aspects where order of elements is not important.

Furthermore, it might be the case that judgments concerning grammatical non-manuals are less reliable than or at least different from judgments concerning manual signs. The motivation for this hypothesis comes from numerous psychological and neurological studies showing that dividing attention is costly (Styles 2008). A signer has to perceive both manual signs and non-manual markers; note that the two are actually in separate parts of the visual field. While non-manuals are clearly linguistically relevant in many cases (recall the discussion of expression of negation), manual signs are indispensable as they contain both lexical and grammatical information. It is thus reasonable to hypothesize that more attention will be allocated to the manual signs, and thus perception of non-manuals will be less focused, leading to less clear acceptability judgments.⁹

In fact, there is some evidence of differences between manual signs and non-manuals in judgment experiments. Davidson (2013) studied scalar implicatures in ASL in two different constructions in a felicity judgment experiment. One condition involved manual signs with quantificational meaning, such as SOME and ALL. It is well known that in spoken languages, quantifiers like *some* create scalar implicatures: a sentence like *Some of the cans are red* would not be used to describe a situation where all the cans are red – even though semantically this sentence is true – due to the presence of a stronger alternative (*All of the cans are red*). Davidson has demonstrated that ASL signers also compute such implicatures and reject a description containing SOME in contexts where ALL could be used, so at least for some manual signs scalar

⁸ Note that these studies show that the difference in memory capacity is not connected to deafness. Crucially, hearing native signers (hearing children of deaf parents who have acquired a sign language from birth) have a larger working memory for words than for signs (Boutla *et al.* 2004).

⁹ It has been shown that signers in fact have enhanced visual attention to the periphery in comparison to non-signers (Bavelier *et al.* 2000); however, this still does not mean that processing both manuals and non-manuals in signed communication is not resource-intensive. It has also been shown that the visual attention is centered mostly at the face of the interlocutor, not at the hands (Neville & Lawson 1987).

implicatures in signed languages work as expected. However, a situation is different when we turn to non-manual markers.

Another example of a scale which produces scalar implicatures in spoken languages is the scale consisting of conjunctions *or* and *and*. While semantically *A or B* is true when both A and B are true, if a speaker of English says *A spoon is in the mug or a spoon is in the bowl*, it implies that it's not true that there is a spoon in the mug and in the bowl (by the same mechanism as with quantifiers). Interestingly, ASL expresses the meanings of 'and' and 'or' by two similar non-manual strategies. The signer leans to one side over the first conjunct, and to the other side over the second conjunct in both contexts; the difference between 'and' and 'or' is expressed by the position of eyebrows. Crucially, it turns out that signers, when seeing a sentence accompanied with disjunctive non-manuals (meaning 'or') in a situation where a conjunctive interpretation is compatible with the context, are much more accepting than with mismatched implicatures of quantifiers. Davidson demonstrates that it is not the case that non-manuals are completely ignored in the judgment: descriptions of disjunctive situations by conjunctive non-manual markers are in fact correctly rejected. Davidson concludes that signers are less likely to compute scalar implicatures for this non-manual marker of coordination, and that the two versions of the marker do not form a scale. This experiment shows that while non-manual marking is used in ASL to disambiguate the two types of conjunction, it does not seem to produce the same semantic/pragmatic effects as manual signs.

I argue that it is thus useful to conduct further investigation on acceptability judgments concerning non-manual markers. Judgments on ungrammatical non-manual markers should be compared to judgments in grammaticality violations purely in the visual channel.

4.3 SOCIO-LINGUISTIC ISSUES

Sign languages are socio-linguistically different from many spoken languages because most signers in the population are non-native (late learners) and bilingual. These facts might influence acceptability judgment experiments in a variety of ways.

Existing research shows that non-native signers are different from native signers in a syntactic acceptability judgment experiment (Boudreault & Mayberry 2006; Cormier *et al.* 2012; Novogrodsky *et al.* 2017). Importantly, the differences are not always trivial: it is not the case that non-native or late learners are always significantly worse in judging all types of structures (Cormier *et al.* 2012) although mostly they do show non-native level of correct answers; see also Chapter 16 of this volume for a discussion of using judgments in L2 speakers.

Most crucially, these studies compared native and non-native signers in only two sign languages (ASL and BSL), with six structures (simple sentences, sentences with negation, sentences with agreeing verbs, wh-questions, relative clauses, and sentences with classifier predicates) all of which were made ungrammatical by changing sign order (recall also the discussion in Section 4.2.2). It might be helpful to conduct more research on judgments in native vs. non-native signers on a larger number of structures and on different types of grammatical violations. The main reason to do such research is practical: if non-native signers turn out to be not different

from native signers for at least some types of tasks, it would make conducting experimental elicitation much easier by making the population of potential participants larger. This would matter most for research on sign languages with a small number of signers (e.g. NGT has 7500 deaf signers including late learners (Wheatley & Pabsch 2012)) for which it is often not realistic to get a sufficiently large sample of native signers.

The fact that most signers are bilingual need not necessarily have an effect on acceptability judgment experiments. For instance, bilingual speakers of spoken languages do not seem to differ in such tasks from monolinguals (Johnson & Newport 1989): speakers are able to suppress one language when they have to make judgments about another one. However, there are reasons to believe that situation with sign-spoken language bilingualism (often called bimodal bilingualism (Quadros *et al.* 2016)) is different – at least in some countries.

The main reason that I think spoken language might interfere more with sign language judgment experiments is the existence of manual communication systems. For instance, RSL coexists with the system which can be called Signed Russian; this system uses signs from RSL but also a large number of artificial signs: for instance, signs for Russian prepositions or even case endings. This system is mostly used in parallel with spoken language and the sign order follows the word order of the spoken utterance. It is often used by hearing people interacting with deaf signers, including sign language interpreters and teachers, but deaf signers also use it among themselves.

Such a situation of coexistence of two means of manual communication – a natural sign language and a partially artificial spoken-language based system – is common to most countries with organized deaf education and interpreting services. For historical reasons (see e.g. Rietveld-van Wingerden 2003), natural sign languages often have a lower status than a spoken language and even the manual communication system. For instance, in Russia, some members of the deaf community assume that more educated deaf people would and should use Signed Russian rather than RSL. This preference for the manual communication system might result in signers accepting spoken-language based structures in an acceptability judgment experiment.

In addition to the issue of lower status of the natural sign language, it is also often difficult to clearly differentiate between the natural sign language and the manual communication system. Signers often switch between the two systems in natural communication. Some signers are not aware of the existence of the two systems, and even of those who are aware, not all are comfortable providing metalinguistic judgments on what belongs to sign language proper and what does not.

Delineation is further complicated by the fact that borrowing of lexical and grammatical elements from the spoken language is a common phenomenon (Adam 2012). For instance, if we find a structure in a corpus of a sign language that looks like a structure from the corresponding spoken language, it might result either from borrowing or from code-switching to the manual communication system, and there is no way to distinguish the two by looking at isolated cases.

All of this implies that, in acceptability judgment experiments, it is not clear whether and how signers would judge structures that result from code-switching or from borrowing. Two undesirable effects can occur: first, signers might accept structures that are grammatical in the spoken language but do not in fact belong to the natural sign language, and second, signers might reject structures that are grammatical in the spoken language and have in fact been borrowed by the sign language or are accidentally similar in the sign language. The former mistake would be caused by the failure to distinguish the sign language from the manual communication system or by a preference for the manual communication system. The latter mistake would be due to hypercorrection: trying too hard to distinguish the natural sign language from the spoken language. It is likely that the formulation of the judgment task might play a role: whether the signers are instructed to distinguishing the natural sign language from spoken language or not; however, given that mistakes can go in both directions, it is not a priori clear how to formulate the instructions properly.

It is thus necessary to empirically test whether judgments on structures which overlap with similar structures in a spoken language are different from judgments on structures which are unique to the sign language. Effects of the formulation of the task should also be studied.

There are some other socio-linguistic factors that should be also taken into account when designing acceptability experiments, such as age, regional variation, and other socio-linguistic variables that are also relevant for spoken languages.¹⁰

Considering age differences, sign languages are comparatively young; some sign languages, such as Nicaraguan Sign Language and Al-Sayyid Bedouin Sign Language have emerged just several decennia ago (Sandler *et al.* 2005; Senghas *et al.* 2004). Research has shown that sign languages change extremely rapidly in the first decennia after their emergence, so signers from different generations in fact have very different grammars (*ibid.*). More established sign languages, such as ASL or BSL, have also been shown to change relatively rapidly (Stamp *et al.* 2015), which means that generational differences can be visible in acceptability experiments.¹¹ While I do not think that it is necessary to conduct acceptability experiments to directly study age effects, such effects need to be taken into account when designing experiments (see also the next section).

Another factor is possible regional variation. Comparatively little research has been done on regional variation in grammar of sign languages, but it is obvious that such variation exists, and for some sign languages, one can speak of the existence of dialects (Schembri & Johnston 2012). Again, this should be taken into account when designing experiments, either by focusing on one regional variant, or by including region as a variable deliberately.

¹⁰ I thank William Matchin (p.c.) for bringing this to my attention.

¹¹ Note that most existing research on age differences in established sign languages concerns lexical or phonological variation, so there is no solid basis to argue that grammar also varies across age groups. However, informal impressions of researchers working in the field suggest that grammatical variation is present.

5. PRACTICAL RECOMMENDATIONS

While in the previous section I have argued that much methodological research is needed before we gain a good understanding of the validity and reliability of acceptability judgment experiments in sign linguistics, some general practical recommendations for researchers using such tasks at the moment can be formulated based on spoken language research and past experiences of sign linguists. I briefly summarize them here.

First, the general methodological recommendations that apply to spoken languages can also be used in creating judgments experiments for sign languages (Schütze & Sprouse 2014), see also various chapters in the first part of this volume. It is important to collect data from multiple signers, to use multiple lexicalizations of the same phenomenon, to pseudorandomize the order of items, to use fillers, and to formulate instructions carefully (see below). The data should be analyzed statistically (see Chapter 3 of this volume). Some sociolinguistic factors can influence experiments in both spoken and signed languages, so possible regional and age variation should be taken into account.

Second, it is important to use the appropriate medium of presentations. Sentences which are to be assessed have to be presented as videos, and so should be the instructions. To minimize possible interference of spoken language, ideally, no spoken language (also in a written form) should be used at any point during the experiment. In addition, it is well known that mere presence of a hearing researcher might increase the influence of a spoken language (Van Herreweghe & Vermeerbergen 2012), so the task should be either administered by a deaf native signer, or be fully computerized so that no communication with the researcher is necessary. Moreover, including deaf signers as collaborators and not only as participants addresses important ethical issues of Deaf identity and language ownership (see Chapters 1-3 in Orfanidou et al. (2015)).

Third, it is very important to formulate instructions in a judgment task appropriately. The instruction for a syntactic task should make it clear that the researchers are not interested in semantic plausibility. A common method in spoken linguistics, which will probably work for sign languages as well, is to ask the participant the following: “If someone signs this sentence, would they look like a native signer?” Although no research on this issue exists yet, it seems reasonable to ask signers to also pay attention to non-manuals if non-manuals are relevant for the phenomenon under investigation.

Neidle et al. (2000) also argue that signers should always be asked to repeat the sentence themselves before judging its acceptability. They claim that signers are often too accepting when they simply view the item and that this effect can be mitigated by asking them to repeat it. While this issue has not been formally investigated, this advice seems reasonable and easy to apply, at least if elicitation is conducted by a researcher and not via the Internet.

As an illustration, I briefly sketch a hypothetical experiment to study the effect of verb class on word order in RSL.¹² In many sign languages, plain (non-agreeing) verbs favor the SVO order, while agreeing verbs favor the SOV order (Napoli & Sutton-Spence 2014). Kimmelman (2012) argued that RSL does not conform to this pattern, as both plain and agreeing verbs favor the SVO order. However, this study was based on a very small-scale dataset of narratives and picture descriptions, so it might be reasonable to revisit this issue.

Following the general methodological considerations, 12 plain verbs and 12 agreeing verbs are selected to avoid the individual effects of specific lexemes.¹³ For each verb, two stimuli are created: one with the SVO order and one with the SOV order, illustrated by (8).

- (8)
- a. BOY INDEX_a aSEE_b GIRL INDEX_b → agreeing verb, SVO
 - b. BOY INDEX_a GIRL INDEX_b aSEE_b → agreeing verb, SOV
 - c. BOY INDEX_a LOVE GIRL INDEX_b → plain verb, SVO
 - d. BOY INDEX_a GIRL INDEX_b LOVE → plain verb, SOV

The acceptability task will contain 12 test items (6 with plain verbs – 3 with each word order, 6 with agreeing verbs – 3 with each word order, such that one participant never sees the same verb twice). Several versions of the task are created: two versions are necessary to include all the verbs in both word orders, and each of these versions should be presented in at least two different item orders. In addition to the test items, the task should contain at least 12 filler items, and 6 training items to be placed in the beginning of the task to get the participant familiar with the procedure. The fillers and the training items target a phenomenon not related to word order.

The test items, the filler items, and the instruction are signed by a native signer of RSL and recorded as video files. The instruction is roughly as follows: “You will now watch 30 short clips of phrases in RSL. Please see whether they look correct or not. To make your assessment, imagine this: if someone signs this sentence, do you think this person is a native signer? Please give each sentence a grade from 0 to 5, where 0 means the signer is clearly not native, and 5 means that the signer is clearly native.” Alternatively, if the experiment will not be conducted via the Internet but by a researcher in person, the instruction can also contain the following: “When you see the sentence, please repeat it first, and then give your grade”.

Next, 12 signers of RSL are recruited as participants. Given that the research question does not concern sociolinguistic variables, only native signers from the same region and from roughly the same age group are selected. Different versions of the task are distributed evenly across

¹² This is a toy example, so only one factor influencing word order is considered, while in reality there are multiple factors that can influence it, and they should be studied together. Furthermore, this example does not address any of the modality-specific or sociolinguistic issues discussed in Section 4.

¹³ The numbers of items and signers given here are mere suggestions; to determine the real number of items and signers necessary, potential effect sizes should be estimated and a power study should be conducted first (Schütze & Sprouse 2014).

participants. The task is presented on a laptop/desktop so that the participants can watch the videos of the instruction and the test items and click on buttons to grade the videos.

The resulting dataset is analyzed using a mixed-effects model (Baayen *et al.* 2008) so that individual effects of participants and lexical items could be taken into account. If we find a significant difference in acceptability of the SVO vs. SOV word order for plain vs. agreeing verbs, we have found evidence against Kimmelman's (2012) claim. If no significant difference is found, the experiment is not informative.

APPENDIX A: GLOSSING CONVENTIONS

Glossing conventions: Signs are glossed with approximate translations in SMALL CAPS. INDEX stands for a pointing sign; subscripts represent a person (1) or a location in space (a, b): pointing signs and verbs can be modified to agree with arguments this way; CL stands for a classifier handshape followed by the form it refers to. Non-manual markers are placed above the glosses, and the underscore reflects the scope of the non-manual.

Abbreviations: ASL – American Sign Language; BSL – British Sign Language; LIS – Italian Sign Language; NGT – Sign Language of the Netherlands; RSL – Russian Sign Language.

REFERENCES:

- Abner, N. (2013). Gettin' together a posse: The primacy of predication in ASL possessives. *Sign Language & Linguistics*, **16**(2), 125–156.
- Adam, R. (2012). Language contact and borrowing. In R. Pfau, M. Steinbach, & B. Woll, eds., *Sign language: An international handbook*, Berlin: Mouton de Gruyter, pp. 841–862.
- Arendsen, J., Doorn, A. J. van, & Ridder, H. de. (2010). Acceptability of sign manipulations. *Sign Language & Linguistics*, **13**(2), 101–155.
- Aronoff, M., Meir, I., & Sandler, W. (2005). The paradox of sign language morphology. *Language*, **81**(2), 301–344.
- Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language*, **59**(4), 390–412.
- Bavelier, D., Tomann, A., Hutton, C., ... Neville, H. (2000). Visual attention to the periphery is enhanced in congenitally deaf individuals. *The Journal of Neuroscience*, **20**(17), RC93–RC93.
- Börstell, C. (2017). *Object marking in the signed modality. Verbal and nominal strategies in Swedish Sign Language and other sign languages* (Doctoral dissertation), Stockholm University, Stockholm.
- Bouchard, D., & Dubuisson, C. (1995). Grammar, Order & Position of Wh-Signs in Quebec Sign Language. *Sign Language Studies*, **1087**(1), 99–139.
- Boudreault, P., & Mayberry, R. I. (2006). Grammatical processing in American Sign Language: Age of first-language acquisition effects in relation to syntactic structure. *Language and Cognitive Processes*, **21**(5), 608–635.
- Boutla, M., Supalla, T., Newport, E. L., & Bavelier, D. (2004). Short-term memory span: insights from sign language. *Nature Neuroscience*, **7**(9), 997–1002.

- Branchini, C., Cardinaletti, A., Cecchetto, C., Donati, C., & Geraci, C. (2013). Wh-duplication in Italian Sign Language (LIS). *Sign Language & Linguistics*, **16**(2), 157–188.
- Burkova, S. (2015). Russian Sign Language Corpus. Retrieved April 1, 2018, from <http://rsl.nstu.ru/>
- Caponigro, I., & Davidson, K. (2011). Ask, and tell as well: Question–Answer Clauses in American Sign Language. *Natural Language Semantics*, **19**(4), 323–371.
- Cecchetto, C. (2012). Sentence types. In R. Pfau, M. Steinbach, & B. Woll, eds., *Sign language: An international handbook*, Berlin: De Gruyter Mouton, pp. 292–315.
- Cecchetto, C., Geraci, C., & Zucchi, S. (2009). Another way to mark syntactic dependencies: The case for right-peripheral specifiers in sign languages. *Language*, **85**(2), 278–320.
- Cormier, K., Schembri, A., Vinson, D., & Orfanidou, E. (2012). First language acquisition differs from second language acquisition in prelingually deaf signers: Evidence from sensitivity to grammaticality judgement in British Sign Language. *Cognition*, **124**(1), 50–65.
- Couvee, S., & Pfau, R. (2018). Structure and Grammaticalization of Serial Verb Constructions in Sign Language of the Netherlands—A Corpus-Based Study. *Frontiers in Psychology*, **9**. doi:10.3389/fpsyg.2018.00993
- Crasborn, O., Zwitserlood, I., & Ros, J. (2008). Corpus NGT. An open access digital corpus of movies with annotations of Sign Language of the Netherlands. Retrieved from <http://www.ru.nl/corpusngtuk/introduction/welcome/>
- Davidson, K. (2013). “And” or “or”: General use coordination in ASL. *Semantics and Pragmatics*, **6**. doi:10.3765/sp.6.4
- Davidson, K. (2014). Scalar implicatures in a signed language. *Sign Language & Linguistics*, **17**(1), 1–19.
- Davidson, K. (2015). Quotation, demonstration, and iconicity. *Linguistics and Philosophy*, **38**(6), 477–520.
- de Beuzeville, L., Johnston, T., & Schembri, A. C. (2009). The use of space with indicating verbs in Auslan: A corpus-based investigation. *Sign Language & Linguistics*, **12**(1), 53–82.
- Divjak, D. (2017). The Role of Lexical Frequency in the Acceptability of Syntactic Variants: Evidence From *that*- Clauses in Polish. *Cognitive Science*, **41**(2), 354–382.
- Emmorey, K., & Herzig, M. (2003). Categorical versus gradient properties of classifier constructions in ASL. In K. Emmorey, ed., *Perspectives on Classifier Constructions in Signed Languages*, Mahwah, NJ: Lawrence Erlbaum Associates, pp. 222–246.
- Featherston, S. (2007). Data in generative grammar: The stick and the carrot. *Theoretical Linguistics*, **33**(3). doi:10.1515/TL.2007.020
- Fenlon, J., Schembri, A., & Cormier, K. (2018). Modification of indicating verbs in British Sign Language: A corpus-based study. *Language*, **94**(1), 84–118.
- Geraci, C., Battaglia, K., Cardinaletti, A., ... Mereghetti, E. (2011). The LIS Corpus Project: A Discussion of Sociolinguistic Variation in the Lexicon. *Sign Language Studies*, **11**(4), 528–574.
- Geraci, C., Gozzi, M., Papagno, C., & Cecchetto, C. (2008). How grammar can cope with limited short-term memory: Simultaneity and seriality in sign languages. *Cognition*, **106**(2), 780–804.
- Gökgöz, K. (2013). *The Nature of Object Marking in American Sign Language* (Doctoral dissertation), Purdue University, West Lafayette.
- Henner, J., Caldwell-Harris, C. L., Novogrodsky, R., & Hoffmeister, R. (2016). American Sign Language syntax and analogical reasoning skills are influenced by early acquisition and

- age of entry to signing schools for the deaf. *Frontiers in Psychology*, **07**. doi:10.3389/fpsyg.2016.01982
- Hoza, J., Neidle, C., MacLaughlin, D., Kegl, J., & Bahan, B. (1997). A unified syntactic account of rhetorical questions in American Sign Language. In C. Neidle, D. MacLaughlin, & R. G. Lee, eds., *Syntactic structure and discourse function: an examination of two constructions in American Sign Language*, Boston: ASLLRP publications, pp. 1–23.
- Johnson, J. S., & Newport, E. L. (1989). Critical period effects in second language learning: The influence of maturational state on the acquisition of English as a second language. *Cognitive Psychology*, **21**(1), 60–99.
- Kayne, R. S. (1994). *The antisymmetry of syntax*, Cambridge: MIT Press.
- Kimmelman, V. (2012). *Word Order in Russian Sign Language*, Vol. 12.
- Kimmelman, V. (2018). Impersonal reference in Russian Sign Language. *Sign Language & Linguistics*, **21**(2), 204–231.
- Kimmelman, V., Klomp, U., & Oomen, M. (2018). Where methods meet: Combining corpus data and elicitation in sign language research. In *Proceedings of 8th Workshop on the representation and processing of Sign Languages: Involving the linguistic community*, Paris: ELRA, pp. 95–100.
- Krebs, J., Wilbur, R. B., & Roehm, D. (2017). Two agreement markers in Austrian Sign Language (ÖGS). *Sign Language & Linguistics*, **20**(1), 27–54.
- Kuhn, J., & Aristodemo, V. (2017). Pluractionality, iconicity, and scope in French Sign Language. *Semantics and Pragmatics*, **10**(6). doi:10.3765/sp.10.6
- Langsford, S., Perfors, A., Hendrickson, A. T., Kennedy, L. A., & Navarro, D. J. (2018). Quantifying sentence acceptability measures: Reliability, bias, and variability. *Glossa: A Journal of General Linguistics*, **3**(1:37), 1–34.
- Liddell, S. K. (2003). *Grammar, gesture, and meaning in American Sign Language*, Cambridge: Cambridge University Press.
- Lillo-Martin, D., & Meier, R. P. (2011). On the linguistic status of ‘agreement’ in sign languages. *Theoretical Linguistics*, **37**(3–4). doi:10.1515/thli.2011.009
- Loos, C. (2017). *The syntax and semantics of resultative constructions in Deutsche Gebärdensprache (DGS) and American Sign Language (ASL)* (Doctoral dissertation), The University of Texas at Austin, Austin.
- Meier, R. P. (2012). Language and modality. In R. Pfau, M. Steinbach, & B. Woll, eds., *Sign language: An international handbook*, Berlin: De Gruyter Mouton, pp. 77–112.
- Mitchell, R. E., & Karchmer, M. A. (2004). Chasing the mythical ten percent: Parental hearing status of deaf and hard of hearing students in the united states. *Sign Language Studies*, **4**(2), 138–163.
- Myers, J. (2009). The design and analysis of small-scale syntactic judgment experiments. *Lingua*, **119**(3), 425–444.
- Napoli, D. J., & Sutton-Spence, R. (2014). Order of the major constituents in sign languages: implications for all language. *Frontiers in Psychology*, **5**. doi:10.3389/fpsyg.2014.00376
- Neidle, C., Kegl, J., Bahan, B., Aarons, D., & MacLaughlin, D. (1997). Rightward wh-movement in American Sign Language. In D. Beerman, D. LeBlanc, & H. Van Riemsdijk, eds., *Rightward movement*, Philadelphia: John Benjamins, pp. 247–278.
- Neidle, C., Kegl, J., MacLaughlin, D., Bahan, B., & Lee, R. G. (2000). *The syntax of American Sign Language: functional categories and hierarchical structure*, Cambridge: MIT Press.

- Neidle, C., MacLaughlin, D., Lee, R. G., Bahan, B., & Kegl, J. (1998). The Rightward Analysis of wh-Movement in ASL: A Reply to Petronio and Lillo-Martin. *Language*, **74**(4), 819–831.
- Neville, H., & Lawson, D. S. (1987). Attention to central and peripheral visual space in a movement detection task: An event-related potential and behavioral study (Parts I, II, III). *Brain Research*, **405**, 253–294.
- Novogrodsky, R., Henner, J., Caldwell-Harris, C., & Hoffmeister, R. (2017). The development of sensitivity to grammatical violations in American Sign Language: native versus nonnative signers: sensitivity to grammatical violations in ASL. *Language Learning*, **67**(4), 791–818.
- Oomen, M., & Pfau, R. (2017). Signing not (or not): A typological perspective on standard negation in Sign Language of the Netherlands. *Linguistic Typology*, **21**(1), 1–51.
- Orfanidou, E., Woll, B., & Morgan, G. (Eds.). (2015). *Research methods in sign language studies: a practical guide*, Hoboken: Wiley-Blackwell.
- Perniss, P., Thompson, R. L., & Vigliocco, G. (2010). Iconicity as a General Property of Language: Evidence from Spoken and Signed Languages. *Frontiers in Psychology*, **1**. doi:10.3389/fpsyg.2010.00227
- Petronio, K., & Lillo-Martin, D. (1997). WH-Movement and the Position of Spec-CP: Evidence from American Sign Language. *Language*, **73**(1), 18–57.
- Pfau, R., & Quer, J. (2010). Nonmanuals: their prosodic and grammatical roles. In D. Brentari, ed., *Sign Languages*, Cambridge: Cambridge University Press, pp. 381–402.
- Pfau, R., Steinbach, M., & Woll, B. (Eds.). (2012). *Sign language. An international handbook*, Berlin: De Gruyter Mouton.
- Quadros, R. M. de, Lillo-Martin, D., & Chen Pichler, D. (2016). Bimodal bilingualism: Sign language and spoken language. In M. Marschark & P. E. Spencer, eds., *The Oxford handbook of deaf studies in language*, Oxford: Oxford University Press, pp. 181–196.
- Rietveld-van Wingerden, M. (2003). Educating the deaf in The Netherlands: a methodological controversy in historical perspective. *History of Education*, **32**(4), 401–416.
- Sandler, W., & Lillo-Martin, D. C. (2006). *Sign language and linguistic universals*, Cambridge: Cambridge University Press.
- Sandler, W., Meir, I., Padden, C., & Aronoff, M. (2005). From The Cover: The emergence of grammar: Systematic structure in a new language. *Proceedings of the National Academy of Sciences*, **102**(7), 2661–2665.
- Schembri, A., & Johnston, T. (2012). Sociolinguistic aspects of variation and change. In R. Pfau, M. Steinbach, & B. Woll, eds., *Sign language: An international handbook*, Berlin: Mouton de Gruyter, pp. 788–816.
- Schlenker, P. (2014). Iconic features. *Natural Language Semantics*, **22**(4), 299–356.
- Schlenker, P. (2017). Super monsters I: Attitude and Action Role Shift in sign language. *Semantics and Pragmatics*, **10**(9). doi:10.3765/sp.10.9
- Schütze, C. T., & Sprouse, J. (2014). Judgment data. In D. Sharma & R. Podesva, eds., *Research Methods in Linguistics*, Cambridge: Cambridge University Press, pp. 27–50.
- Sehyr, Z. S., & Cormier, K. (2015). Perceptual categorization of handling handshapes in British Sign Language. *Language and Cognition*, 1–32.
- Senghas, A., Kita, S., & Özyürek, A. (2004). Children Creating Core Properties of Language: Evidence from an Emerging Sign Language in Nicaragua. *Science*, **305**(5691), 1779–1782.
- Sprouse, J., Schütze, C. T., & Almeida, D. (2013). A comparison of informal and formal acceptability judgments using a random sample from Linguistic Inquiry 2001–2010. *Lingua*, **134**, 219–248.

- Stamp, R., Schembri, A., Fenlon, J., & Rentelis, R. (2015). Sociolinguistic variation and change in British Sign Language number signs: Evidence of leveling? *Sign Language Studies*, **15**(2), 151–181.
- Styles, E. A. (2008). *The psychology of attention*, Hove: Psychology Press.
- Taub, S. F. (2012). Iconicity and metaphor. In R. Pfau, M. Steinbach, & B. Woll, eds., *Sign language: An international handbook*, Berlin: De Gruyter Mouton, pp. 388–412.
- Thompson, R. L., Vinson, D. P., Woll, B., & Vigliocco, G. (2012). The road to language learning is iconic: Evidence from British Sign Language. *Psychological Science*, **23**(12), 1443–1448.
- Van Gijn, I. (2004). *The quest for syntactic dependency. Sentential complementation in Sign Language of the Netherlands*. (Doctoral dissertation), University of Amsterdam, Amsterdam.
- Van Herreweghe, M., & Vermeerbergen, M. (2012). Handling sign language data. In R. Pfau, M. Steinbach, & B. Woll, eds., *Sign language: An international handbook*, Berlin: De Gruyter Mouton, pp. 1023–1045.
- Wheatley, M., & Pabsch, A. (2012). *Sign language legislation in the European Union*, 2nd edn, Brussels: European Union of the Deaf.
- Wilbur, R. (1996). Evidence for function and structure of wh-clefts in American Sign Language. In W. H. Edmondson & R. Wilbur, eds., *International Review of Sign Linguistics*, Erlbaum, pp. 209–256.
- Wilbur, R. B. (2008). Complex predicates involving events, time, and aspect: Is this why sign languages look so similar? In J. Quer, ed., *Signs of the time. Selected papers from TISLR 2004*, Hamburg: Signum, pp. 217–250.
- Wilbur, R. B., & Schick, B. S. (1987). The effects of linguistic stress on ASL signs. *Language and Speech*, **30**(4), 301–323.
- Wilson, M., & Emmorey, K. (1997). A visuospatial “phonological loop” in working memory: Evidence from American Sign Language. *Memory & Cognition*, **25**(3), 313–320.
- Woodward, J. (1994). *Describing variation in American Sign Language: Implicational lects on the deaf diglossic continuum*, Burtonsville, MD: Linstok Press.
- Zwitserslood, I. (2012). Classifiers. In R. Pfau, M. Steinbach, & B. Woll, eds., *Sign language: An international handbook*, Berlin: Signs of the time. Selected papers from TISLR 2004, pp. 158–186.

LIST OF FIGURES

Figure 22.1 RSL sign ₁CL(ROUND)-GIVE_a (first and last frames)