Radboud University

The contribution of lexical overlap

to perceived iconicity in foreign signs

Author: Rehana Omardeen

Student number: 4683277

First supervisor and first reader: Prof. dr. O. Crasborn

Second supervisor and second reader: dr. C. Börstell

September, 2018

Acknowledgements

I would like to say thank you my supervisors, Onno Crasborn and Calle Börstell for their advice and guidance throughout the thesis project. Special thanks to Merel van Zuilen for recording the CSL signs, and Roeland van Hout for help with mixed models. Lastly, thank you to everyone who helped recruit participants for the study, and of course the participants themselves. The handshape fonts used in this thesis are created by the Centre for Sign Linguistics and Deaf Studies, CUHK.

List of Abbreviations

ASL	American Sign Language	
BSL	British Sign Language	
CSL	Chinese Sign Language	
DGS	German Sign Language	Deutsche Gebärdensprache
Libras	Brazilian Sign Language	Língua Brasileira de Sinais
LSC	Catalan Sign Language	Llengua de Signes Catalana
NGT	Sign Language of the Netherlands	Nederlandse Gebarentaal
TİD	Turkish Sign Language	Türk İşaret Dili

1. Introduction

Sign languages like spoken languages are natural languages that can be analysed at all levels of linguistic complexity. However, they have not always been recognised as such. The iconic properties of these languages have led people to believe that these languages are holistic and pantomimic, lacking internal structure. In fact, the case is quite the opposite. Early on, linguists have identified that signs, like words have phonology and can be broken down to their sub-lexical structure. The pioneering work of Stokoe (1960) and Battison (1973) identified four major formational units of a sign; handshape, location, movement, and orientation. These components of a sign are critical to its well-formedness and can account for the minimal difference in signs that are otherwise identical. The example below, taken from Sign Language of the Netherlands (Nederlandse Gebarentaal; NGT), shows one such minimal pair of signs. The signs ZEGGEN (*say*) and BESTELLEN (*order*) are distinguished by movement but identical in all other parameters.[1]

Figure 1. Minimal pair ZEGGEN (1a) and BESTELLEN (1b), distinguished by orientation in NGT

More recently, models of sign phonology have transitioned to the use of hierarchical auto-segmental models to describe sign form. These models have outlined more complex and compositional structures embedded within traditional formational parameters (e.g. Brentari, 1998; Sandler, 1989; van der Hulst, 1993). What were once considered basic formational units can be further dismantled. For example, instead of Stokoe's parameter of *orientation*, a sign can be analysed for whether it changes orientation (*orientation change*), or what the orientation of the sign is relative to its movement path or location (also known as its *relative orientation*; Crasborn & van der Kooij, 1997). Returning to the example from NGT (Figure 1), we can pinpoint the differences between the sign ZEGGEN and the sign BESTELLEN to a difference in the orientation of the finger relative to movement path. In fact the two signs align in other aspects of orientation.

---

[1] Following convention, signs will be glossed using small caps. To signify that a sign is from a particular language, the abbreviation for the language will precede the sign gloss, e.g. NGT-ZEGGEN.

Understanding the phonology of signs has not only given us a window into sub-lexical structure, but has also given us a frame to map the iconicity observed in sign languages at the level of the sign. Various works have examined in detail how meaning is encoded at the parameter and feature level, with linguists offering ways of integrating iconicity into phonological models (e.g. Eccarius & Brentari, 2010; van der Kooij, 2002). In fact, the rise of sign language linguistics has provided key evidence that challenges traditional views on iconicity. Foundational views of linguistics posit language to be arbitrary, privileging the purely conventional and symbolic connection between form and meaning (de Saussure, 1916). However, sign languages exploit visual mappings between form and meaning at various levels of linguistic structure, resulting in language rich in iconic expression. The existence of such robust iconicity in natural languages has forced linguists to not only re-evaluate the status of iconicity in language, but to explore the patterns, structure and consequences of this iconicity.[2]

These iconic mappings, facilitated by the visuospatial modality, appear to underlie lexical similarities even across unrelated sign languages. Take, for example, the sign FOOD in Turkish Sign Language (Türk İşaret Dili; TİD) and American Sign Language (ASL). These two languages with no known relationship both have identical signs for the concept *food* depicting the hand moving towards the mouth. Such observations have fed into many misconceptions about sign languages, such as the myth that all sign languages are essentially same. This of course, is easily disproven by expanding the comparison of ASL and TİD to other signs, demonstrating that the sign language lexicons are subject to rich variation. Another widely held belief is that iconicity is an objective property of signs. In fact, this too is false, and recent research has begun to illustrate that several aspects of an individual's experience can affect how they perceive iconicity of a given sign.

For copyright purposes, this image has been removed

Figure 2. TİD-EAT (2a) and ASL-EAT (2b) from the Spread the Sign online dictionary[3]

One interesting source of variation between individuals is their native language experience. Indeed, given the rising interest *cross-signing*, the communication between deaf people without a shared sign language (Bradford et al. 2013), examining sensitivity to iconicity across language barriers may provide key insight into what feeds successful communication. Recent research has shown that signers perceive lexical signs from their own language to be more iconic than translation equivalents from a foreign sign language (Adam et al., 2007; Occhino et al., 2017). This suggests that signers' experience using their native language guides their perception of iconicity in foreign signs. Framing this finding within cognitive frameworks of iconicity,

---

[2] This has resulted in a general shift in attitudes, with iconicity being widely recognised today as a core property of language in general, both spoken and signed (Perniss et al., 2010; Dingemanse et al., 2015).

[3] Source: http://www.spreadthesign.com/ (accessed September 1, 2018).

this native language sensitivity can be constructed by the repeated use of form–meaning mappings of the native lexicon. In short, it is possible that signers' perceptions of iconicity are mediated by their native lexical knowledge. Where foreign signs differ from native translation equivalents in these form–meaning mappings, signers may construe these forms as less iconic. This observation provides interesting grounds for further investigation. Specifically, when judging iconicity in a foreign sign, are signers influenced by familiar form–meaning mappings present in the native lexicon? How can we measure this familiarity? While perceived similarity by signers offers one source of insight, can we also capture familiar mappings by examining a more fine-grained measure of form-similarity among signs?

Outline of the thesis

In Section 1 I have laid out a preliminary motivation for the research question. Section 2 will go onto give the necessary background to motivate the study. Section 2.1 focuses on the notion of iconicity in sign languages, first examining how it is embedded in the lexicon, then exploring how perceptions of iconicity can differ across individuals. I then go on to describe studies of lexical comparisons in Section 2.2, examining the methodological drawbacks and benefits of previous work. Section 2.3 motivates the present study and introduces the specific research questions to be addressed. Section 3 will discuss methods, and Section 4 will outline the results of the analysis. Section 5 will first summarise these results, then discuss some methodological challenges, and finally places the results in a broader context, offering angles for future research. Section 6 concludes.

2. Background

2.1. Iconicity

For this study, I adopt the definition of iconicity as put forth by cognitive linguists. Thus, iconicity can be defined as the structured mapping between meaning and form. This process can be thought of as one of analogue building between semantic and phonological representations (Taub, 2001). The model below exemplifies this process with the ASL sign TREE. detailing the steps between semantic and phonological poles of iconic mappings. Starting with meaning, the concept of *tree*, first (1), a sensory image is selected and undergoes a schematisation process (2) where is highlighted which aspects of it are to be mapped. Lastly, (3) the process of encoding fits these schemas onto linguistic form by mapping them onto the articulators.
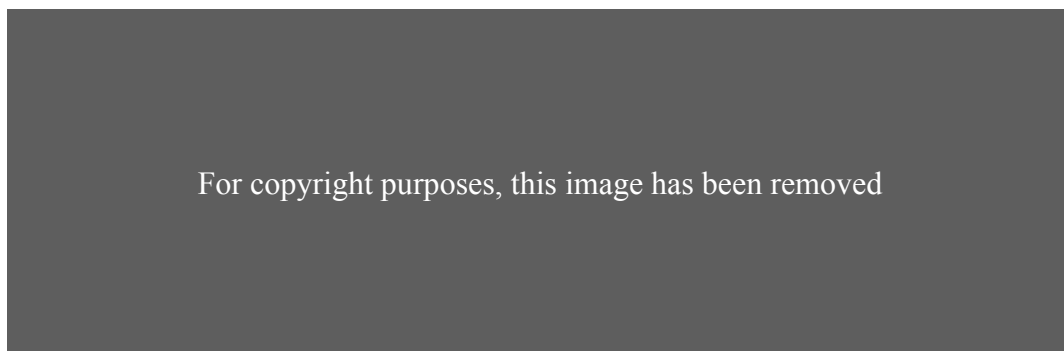

For copyright purposes, this image has been removed

Figure 3. Analogue-building process for ASL TREE (Taub, 2000; p.35)

This process of analogue-building can also extend to metaphorical concepts. Thus, a concept such as *future* may extend to an image through the metaphor FUTURE IS AHEAD. The image of forward motion is then selected, schematised and mapped to the articulators in much the same way as Figure 3, producing a sign that may move forward in space to reflect moving forward in time (Taub, 2001). In either case, whether abstract or concrete, the process of analogue-building is one that creates structure-preserving form–meaning mappings (Meir, 2010; Emmorey, 2014).

Iconicity and sign structure

One key insight that the study of sign phonology has provided is that iconicity can be encoded within the structural components of a sign. In fact, examining the phonological structure of signs at a fine grained level may be critical to properly cataloguing iconic mappings. As Wilcox states, to fully understand iconicity, "it is necessary to describe […] specific handshapes and their features, specific movements with associated manners, paths and so forth, in order to discover their similarity to semantic structure" (Wilcox, 2004; p. 1245). Returning to the example of TREE in Figure 3, the handshape of the dominant hand maps to the branches of the tree, the orientation of the non-dominant hand maps to the ground below the tree, the relationship between the two articulators represents the perpendicular relationship between tree and ground. While iconicity can permeate all features of signs such as ASL-TREE, other signs may contain only some iconic mappings. Thus, iconicity is not holistic but rather compositional.

Early on, sign language phonologists have identified iconicity in the building blocks of signs. Alongside meaningless phonological units, phonologists have identified iconic meaning-bearing and therefore

morphemic elements (e.g. Mandel, 1977; Boyes-Braem, 1981; Brennan, 1990). These findings have challenged a core property of language, namely duality of patterning, the building of meaningful units, words and morphemes, from meaningless sub-structures, phonemes (Hockett, 1960). While traditional phonemes have a purely symbolic and arbitrary connection between form and meaning, instead sign language sub-units have the potential to carry motivated, non-arbitrary mappings between form and meaning.

These mappings have been noted to demonstrate robust patterning both within and across sign language lexicons. Focusing within a language, we can examine the NGT signs LOOK, READ, SEARCH and SEE, each of which employs the V handshape[4]. In these signs, the selection of the handshape parameter is not merely arbitrary. Instead, the two fingers are mapped to the gaze of the two eyes, in a mapping that recurs across various signs in the lexicon. These mappings can also recur across languages, with modality induced iconic similarities attested unrelated sign languages. Returning to the example of FOOD from Section 1, it is not surprising, yet it is notable that in examining the online dictionary Spread the Sign, all 26 signs from various sign languages for the concept *food* are produced at or around the mouth.[5] While some of these languages may be related, others have absolutely no history of contact, yet they still employ the same mapping of the semantic concept *food* to the phonological location mouth.

In a study using this online dictionary database, Östling and colleagues (2018) demonstrate that examining specific concepts across multiple sign languages reveals consistent mappings to specific body locations; some examples found include concepts such as *food* mapping to the mouth, *think* mapping to the head, *love* mapping to the chest. These correspondences can be found in other formational parameters as well. The same study revealed cross-linguistic patterns in the number of hands used to produce a sign (also called *handedness*), echoing findings from previous work. Concepts with specific salient images that depicted interaction between distinct entities (e.g. *meet*), the physical relationship between two entities (e.g. *near*), the physical dimensions of entities (e.g. *house*), or the component parts of entities (e.g. *bicycle*) were more likely to be encoded using two hands across all languages in the sample (Östling et al., 2018; Lepic et al., 2016).

Aside from this large-scale cross linguistic work, other research has compared signs from unrelated language pairs with smaller scale in depth analyses. Occhino (2016) examines the semantic correlates of the feature of *handshape change*, across the lexicons of ASL and Brazilian Sign Language (Língua Brasileira de Sinais: Libras). Looking at signs from both languages, she identifies general semantic schemata that fit to specific handshape changes. In both ASL and Libras she finds that handshape opening can be mapped to concepts that fit within the general schema of *emergence*, including physical emergence with concepts such as *grow*, metaphorical emergence such as the concept *inspire* (Occhino, 2016; p.168). Similarly, she maps handshape closing to *containment* or *grasping* across both languages with signs like *catch* and *accept* fitting both concrete and metaphorical patterns of *grasping* (Occhino, 2016; p.174).

These investigations of cross-linguistic patterns have been so far limited to certain formational parameters, with considerable focus on *handshape* and *location*. There is still much to explore of how other parameters

---

[4] Here, the term V handshape refers to the following hand configuration:

[5] Source: http://www.spreadthesign.com/ (accessed September 1st, 2018).

like *movement* and *orientation* may display similarities in iconic mappings across languages. However, these comparisons illustrate that it is possible to trace across languages the mapping of specific form features to semantically coherent groups of concepts. This is not to say, however that all sign languages obligatorily employ the same mappings. On the contrary, the presence or absence of these mappings can appear quite random. In other cases, they may vary systematically, with patterns of iconicity that display typological variation across sign languages. One such example comes from Padden and colleagues (2013), who find this this systematic variation in the handshapes used to encode tools in diverse sign languages.

Indeed, it is important to remember that there is no one way to map a given concept iconically to a sign form (Klima & Bellugi, 1976). By examining the signs for the same concept in different sign languages we can see how each language may select a different image to encode, resulting in signs that are both iconic but in different ways. For example, while the NGT sign CAT highlights the whiskers of a cat, the Swedish Sign Language sign CAT instead depicts a cat being petted. Which iconic image is selected may even reveal something about a referent that is of particular salience within a culture or community. For example, in an analysis of Catalan Sign Language (Llengua de Signes Catalana; LSC), Jarque (2005) emphasises the importance of sociocultural context in iconic mappings of signs. Her research focuses on abstract metaphorical mappings that are entrenched in the LSC community like FEELINGS ARE AIR or TEACHING IS FEEDING. However, by examining a wider scope of sign languages, it becomes clear that such cultural context may even be important when it comes to encoding highly concrete concepts. For example, in various deaf community sign languages used in urban environments, the sign PIG maps the snout of a pig; whereas in many rural sign languages used in small village communities, the sign pig maps instead the technique used to kill a pig by draining its blood from the neck. Thus, for both abstract metaphorically mapped and highly concrete referents, cultural context may inform which mappings are selected. This kind of cultural context can not only feed these representations, but also feed our interpretation of them, such that an outsider to the rural context may find the iconicity of the sign PIG quite inaccessible. In fact, recent research has confirmed that perceptions of iconicity can differ based on various aspects of individual experience.

Perceived Iconicity

Cognitive frameworks have not only helped connect iconicity to the structural properties of a sign, but have also highlighted important considerations about how these iconic mappings are perceived. Wilcox (2004) situates the form and meaning representations involved in the analogue-building process of iconicity all within an individual's conceptual space. In keeping with ideas from cognitive grammar, he defines *cognitive iconicity* as the distance between the semantic and phonological representations of the symbolic structures that join the two. Thus, some representations are considered more iconic because they more closely resemble each other and are located closer in conceptual space. Critically, this resemblance is not objective, but instead defined by the individual and their ability to create conceptual mappings (Gentner and Markman, 1997).

Thus, while iconicity is often thought to be an objective or inherent property of the sign, cognitive frameworks highlight the fact that much like all other aspects of language that map form to meaning (e.g. grammatical rules), the representations and processes involved in creating iconic mappings are mediated by the individual. Unlike words, which differ from language to language, iconicity has a particularly objective

flavour due to the shared experiences across language communities. The sign FOOD, produced at the mouth, seems objectively iconic because we all eat with our mouths. However, an example such as PIG demonstrates how community level differences in experience may give rise to differences in how iconicity is construed.

In fact, recent research has begun to highlight that one's experiences can affect perceptions of iconicity. Specifically, by looking across language communities, research has examined the role of language experience on iconicity judgements. The bulk of this work thus far has compared signers, those with sign language experience, to non-signers, those without. Early evidence has suggested that experience with a visual language can affect sign transparency, with Pizzuto & Volterra (2000) demonstrating that signers are able to guess the meaning of signs from a foreign sign language more easily than non-signers. More recently, the way non-signers experience iconicity in sign has been linked to gesture. Ortega, et al., (2017) demonstrated that non-signers judge signs that overlap with their gestural repertoire to be more iconic than signs that do not. Much research has also used non-signers as control groups to provide a baseline measure of iconicity, given their lack of linguistic experience.

However, disentangling the effects of linguistic experience on perceiving sign iconicity is a clear challenge, and one that cannot be fully explored by examining how non-signers experience iconicity in a sign language or even how signers experience iconicity in their own language. A study by Adam and colleagues (2007) took a different angle, by examining how deaf signers perceive iconicity in both native and foreign signs. To do so, they presented signers of German Sign Language (Deutsche Gebärdensprache: DGS) with signs from a foreign language, ASL and asked signers to rate the iconicity of foreign and native translation equivalents. The results demonstrated a native language bias in iconicity ratings, DGS signers found signs from their native language to be more iconic than the foreign ASL signs. In a follow up study, Occhino and colleagues (2017) performed a two-way rating task with a matched group of DGS and ASL signers, finding similar results. Native language biases in iconicity ratings persisted across both groups of signers. In fact this bias was observed even in the group of signs rates as most-highly iconic by the foreign language community; these signs were still judged to bee less iconic than their native translation equivalent.

This provides clear evidence that signs are not objectively iconic but instead the experiences of the language user reinforce and shape perceptions of iconicity. When these signs are shared among a community of signers, this may create the illusion of objectivity. However, exposing signers to foreign signs reveals that this perceived iconicity is instead subjective. To explain the native language bias in iconicity ratings, the authors propose that perceived iconicity can be understood as "a construal of form–meaning relationships within a lexical network specific to individual signers" (Occhino et al., 2017; p.109).

These observations open up interesting avenues for further investigation. If language experience influences perceived iconicity, and form–meaning relationships are reinforced by use native lexicon, then how might existing lexical knowledge influence judgements of foreign iconicity? One clue comes from the aforementioned studies, where researchers also asked signers to judge form-similarity between each pair of translation equivalents. They found that even for pairs of signs such as DGS-BANANA and ASL-BANANA (Figure 4) that were rated as highly similar, signers rated native signs as more iconic than foreign. Interestingly however, the iconicity ratings for these highly similar signs appeared to be higher than average.

What this might suggest, is that when viewing foreign sign such as BANANA, which have high overlap with their native lexicon, signers recognise familiar form–meaning mappings and perceive these sign to be very iconic.

For copyright purposes, this image has been removed

Figure 4. DGS-BANANA (4a) and ASL-BANANA (4b) (Adam et al., 2007; p.217)

In summary, iconicity, the structured mapping between form and meaning, can be traced by examining the very building blocks of signs. Sub-lexical features can map to semantic content, and patterns of mappings can occur within and across sign language lexicons. Important to note however is that these sign forms, even when mapped iconically, can be perceived differently based on an individual's experiences. While the iconicity of certain forms may appear objective, specific factors such as native language experience have been shown to affect perceptions of sign iconicity.

2.2. Lexical comparisons

To date, various studies have been interested in investigating similarity between sign translation equivalents from different languages. To assess the relationship between two signs researchers have employed pairwise comparisons, putting these translation equivalents side by side and comparing them on the basis of their formational parameters. Each formational parameter is evaluated among sign pairs and are determined to be either matching or not matching. Then, for each pair, the number of matching parameters is tallied. Where all parameters match, there is full overlap and a sign is judged to be identical. Where one or more parameters do not match, this partial overlap result in signs being judged as either similar or different. The example in Figure 5 below shows a pairwise comparison for the signs NGT-CAT and CSL-CAT, compared across the parameters of *handshape*, *location*, *movement*, *orientation* and *other* (a category included to capture differences that fall outside the traditional four parameters).

Parameter	Match
Handshape	✓
Location	✓
Movement	✓
Orientation	✗
Other	✗

Figure 5. Lexical comparison of NGT-CAT (5a) and CSL-CAT (5b)

This work has been done mainly in the realm of lexicostatistics, where linguists use these comparisons to assess relatedness between languages or language varieties. These studies examine hundreds of sign pairs to identify cross-language cognates, with higher cognate counts leading to higher relatedness. The priority of such studies is to draw conclusions about the lexicons as wholes, as a result, the phonological comparisons are quite rough at the level of the individual sign. This has led to some inconsistencies in methodology that have made the findings of these studies difficult to interpret and compare.

The first methodological challenge is the selection of signs to compare, with some studies employing random lists sourced from sign language dictionaries (e.g. Al-Fityani & Padden, 2010), and others relying on standardised word lists for cross linguistic comparison (e.g. Currie et al., 2002). Interestingly, this choice can affect the results of the comparison, with core vocabulary lists often returning higher rates of cognates than random samples of the lexicon (McKee & Kennedy, 2000). This makes overlap estimates difficult to compare across studies, as a study that compares a random dictionary sample is likely to return less overlap than a study that compares core vocabulary.

The second major methodological concern arises with method of comparison itself. There is little consistency among different studies as to which parameters are compared. While some studies only compare across three parameters of *handshape*, *location* and *movement* (Currie et al., 2002), other studies have added *orientation* (Al-Fityani & Padden, 2010). Others still have added parameters for *number of hands* (Sasaki, 2009), or a catch-all *other* category (McKee & Kennedy, 2000). This leads to differing criteria for the rating of similarity amongst signs. The result is a categorical division between *similar*, *identical* and *different* signs. However, across the aforementioned studies, signs might overlap in any range of 2/3, 3/4, or 3/5 parameters to be counted as similar. Returning to the example of CAT, comparing *handshape*, *location* and *movement*, the two signs may be considered to be identical in line with Currie et al.'s (2002) analysis. Both use identical F hand configurations[6], which begin at the cheek location and move outwards. Comparing *orientation* would result in a mismatch, relegating it to the *similar* category under Al-Fityani & Padden's (2010) coding scheme. Adding McKee et al.'s (2000) *other* category may capture subtle differences between the signs, such as the hand-internal movement of the NGT sign with the closing of the thumb and forefinger[7], or the initial contact of the CSL sign where the fingertips brush the cheek before the sign moves outward. These differences would then render these two signs as *different*.

More recently, studies have begun to refine measures of sign similarity, promising a more detailed estimate of linguistic distance between two signs based on current understanding of sign phonology and a wider range of formational features. In one approach, Yu, Geraci and Abner (2018) propose a new method of lexical comparison that preserves the categorical distinction between *identical*, *similar* and *different* signs but compares at a more detailed level than the formational parameter. They identify that in many previous pairwise comparisons it is unclear what constitutes significant difference between phonemes. This problem is exacerbated by the complexity of formational parameters that can each be broken down into multiple

----

[6] Here, the term F handshape refers to the following hand configuration: .

[7] See https://www.lerengebaren.nl/gebarenwoordenboek/poes (accessed September 1st, 2018).

constituent parts. Indeed, current models of sign phonology reflect this complexity with feature-based descriptions (see Section 2.1).

Yu and colleagues (2018) propose a method of comparison that relies on these feature-based analyses of signs. First, a sign is coded for a range of form-features, these dismantle formational parameters into their building blocks. For example, the parameter of *movement* is captured by a range of features such as *movement direction*, *movement shape*, *movement repetition*, also capturing hand internal movement with features such as *orientation change*. After this coding is complete, signs are compared feature by feature. Signs that are coded the same for all features are labelled as *identical*, signs that differ on one feature are labelled as *similar* and those that differ on more than one feature are labelled as *different*. From these categories, linguistic distance is estimated among language pairs by assessing percentage of cognates.

Such a method also addresses another shortcoming of previous lexical comparisons; there is a fair bit of subjectivity involved in judging two features of a sign as same or not, especially when comparing parameters such as movement with several sub-features. Returning once again to the example of CAT, the handshapes produced in both signs are phonetically identical. In fact, these handshapes, while formationally identical may also be analysed as having different selected fingers, with the handshape in NGT-CAT selecting the thumb and forefinger, and the handshape of CSL-CAT selecting the middle, ring and pinkie finger. Such intricacies may be captured in individually coding signs, but is likely to be lost when comparing parameters side by side. By performing independent phonological coding for each signs, which can then be run through automated comparison programs, as done in this study, the comparisons is shifted away from subjective judgements of similarity and instead there comparison rests on more objective phonological descriptions of each sign.

In another approach, Parks (2011) explores new ways of calculating distance between sign pairs by adapting Levenshtein's distance measurements for sign languages. Levenshtein distance is an algorithm used to calculate the edit distance between two forms; the number of changes necessary to change form A to form B. Typically employed in comparisons of spoken language varieties, Levenshtein's calculations are performed by putting the phonetic representations of two different word forms side by side, then calculating the edit distance. Edits can come in three forms; substitutions, insertions and deletions. In order to make these Levenshtein values comparable across word forms, they are generally normalised. This can be done by aligning forms side by side, then dividing the number of edits by the number of alignment slots for each word pair (Heeringa, 2004). The example below, taken from Beijering et al. (2008), shows the Levenshtein's distance calculated for the form *enige* ("in agreement") between its respective pronunciations in Lyngby (Danish) as [ʔeːni] and Stockholm (Swedish) as [eːnɪɡɑ]. As shown below, 4 changes are made, this number is divided by the total number of 6 alignment slots to leave a normalised Levenshtein's distance 0.67.

For copyright purposes, this image has been removed

Figure 6. Levenshtein's distance calculation for two forms of *einge* (Beijering et al., 2008; p5)

To adapt this method, Parks compares translation equivalents based on 6 parameters: *initial handshape*, *final handshape*, *initial location*, *final location*, *palm orientation change* and *joint movement*. For each sign pair in his comparison, he then calculated how many of these six parameters needed to be changed to get from one sign to the next. Parks aggregates the output of these Levenshtein comparisons across language pairings to determine a measure of linguistic distance between language varieties in his sample. Unlike most previous studies, this method offers a quantitative estimate of linguistic overlap, not just between language pairs, but at the level of the individual sign. This numeric calculation offers possibilities that extend past the previous classification scheme of *identical*, *similar* or *different*, and can allow for more informative insight into the degree to which different signs may overlap.

These new methods highlight two important considerations for determining comparisons between signs of two different languages. Firstly, when considering a comparison based on sub-lexical units of a sign, it is important to incorporate current understanding of sign phonology. Secondly, by expanding the range of form-features from the traditional four parameters, we can escape the categorical distinctions between *similar, identical* and *different*. Instead, it is possible to quantify the distance between any sign pair into a discrete number: an overlap score. This can allow research to not only pinpoint the differences for similar signs, but also create a quantitative degree of how similar two signs are.

2.3. The current study

While previous research has demonstrated a native language bias in perceptions of iconicity in foreign signs, one question that remains unanswered is how language-internal patterns and linguistic experience of signers might affect these perceptions. Indeed studies have connected formational overlap to perceived iconicity, suggesting higher form overlap between translation equivalents leads to higher iconicity ratings in foreign signs. The current study connects lexical comparisons of signs to perceived iconicity scores to examine whether form-overlap with the native lexicon can affect how signers experience iconicity in a foreign sign.

To do so, I propose a new method of lexical comparison, one that compares pairs of translation equivalents at a fine-grained level to produce a quantifiable overlap score between signs. I will connect these scores to an iconicity rating task to see if phonological overlap can predict perceived iconicity in foreign signs. Adapting the methods of Occhino et al., (2017) and Adam et al., (2007), I will use a one-way rating task, in which signers of a given sign language rate signs from their native language and a foreign language. If the native-bias in perceived iconicity can be explained in part by the amount of lexical overlap, then foreign signs that have high form-overlap with the native lexicon should be perceived as more iconic than foreign signs that have low form-overlap with the native lexicon.

Research Questions

To investigate this, I will ask two major research questions.

RQ1	Are signers more likely to judge native signs as more iconic than foreign signs?
RQ2	Does degree of form overlap with native signs help explain how signers rate iconicity in foreign signs?

Based on previous results, I hypothesised that both RQ1 and RQ2 would return positive results. Signers were expected to judge native signs to be more iconic than foreign signs. While using a one-way rating task, these findings will support previous two-way rating studies that demonstrate a native language iconicity bias. In addition, foreign signs that overlap more with the native translation equivalent were expected to elicit higher iconicity ratings than those that do not. As in previous work, I expected to observe this effect through a positive relationship between signers' similarity ratings of translation equivalent pairs and their iconicity ratings of foreign signs. Thus, signs that were rated to be highly similar scores were also expected to be rated as highly iconic. I also expected to see a similar positive relationship between iconicity ratings and a more objective measure of sign similarity, namely the novel measure of phonological overlap. Foreign translation equivalents that had a high phonological overlap score were also expected to be rated as highly iconic.

Languages

This study employs lexical items from two sign languages: Sign Language of the Netherlands (NGT) and Chinese Sign Language (CSL).

NGT is the national sign language used by the deaf community of the Netherlands. It is used by roughly 16,000 people both deaf and hearing and has roots in the French Sign Language of the 19th Century (Crasborn, 2001). There is considerable lexical variation across the various regions of the country, and dictionary and corpus projects have captured this variation with documentation projects in the past few decades (e.g. Schermer et al., 1988; KOMVA, 1989; Crasborn et al., 2008). There has also been considerable linguistic research done on NGT, with the use of lexical databases facilitating work on phonological structure (Crasborn et al., 2001; Crasborn et al., 2015).

The other language used in this study is CSL. There are two major regional varieties of CSL distinguished primarily by lexical variation, the North regional variety and the South regional variety (Yang, 2015). For this study I will use CSL to refer to the Southern variety, specifically the variety used in the city of Shanghai. The exact number of language users of CSL is still undocumented, however 2006 census data estimates over 20 million people in China to be deaf or hard of hearing (Yang, 2015). Linguistic research into CSL is still in its infancy, however recent work has begun to explore the structure of different varieties of the language.

There is no known historical or present day relationship between NGT and CSL, making these two languages an ideal selection to examine the role of phonological overlap without the confounding factors of relatedness or contact. It ensures that there is no etymological link between translation equivalents from different languages, as opposed to related language pairs, where motivations in sign form may been attributed to etymological links. The lack of present day contact between the two language communities further adds the control of exposure and familiarity. Unlike ASL, which is today used on a global scale, CSL is a language which NGT signers are unlikely to have any significant exposure or familiarity with, thus their judgements of iconicity will be based on initial impression of form alone.

3. Methods

3.1. Research design

To answer the research questions, I adopted a two-part design for this study. In the first part, I performed a lexical comparison between translation equivalents from NGT and CSL. This comparison was based on a range of 20 phonological features that were independently coded for each sign, and it resulted in a numeric phonological overlap score for each sign pair. The results from this comparison formed the basis of stimuli for an iconicity rating task, the second part of the study. In this task, signers of NGT viewed and responded to pairs of translation equivalents, producing iconicity ratings and similarity scores for each. The methods for both parts are outlined below.

3.2. Materials

> *Global Signbank*

Signs for this study were sourced from the database Global Signbank (Crasborn et al., 2018b), a lexical database that contains signs from various sign languages. Building off its precursor, NGT Signbank (Crasborn et al., 2018a), this database stores signs clips linked to detailed form-based information, as well as information on morphology, semantics and for NGT, corpus frequency. Signs from each language within the database are coded according to a single detailed phonological coding scheme (as outlined in Section 3.3.2 below). For this study, a subset of Signbank entries from the NGT and CSL datasets were used in the phonological comparison and subsequent iconicity rating task.

The NGT dataset comprises lexical items that were gathered from the annotation project of the Corpus NGT at Radboud University. As these signs are collected from corpus annotations, this dataset contains a great deal of lexical variation that reflects regional varieties of NGT across the Netherlands. These variants are encoded in Signbank as separate entries connected by ID-gloss. Thus, three separate lexical variants for the sign CAT would be stored as separate entries listed as CAT-A, CAT-B and CAT-C. Each NGT sign entry contains the sign, the English gloss, the Dutch gloss and the different possible translation equivalents in Dutch.

The CSL dataset is comprised of signs from Chinese Sign Language collected from the signing community of Shanghai. Signs in this dataset have been collected by researchers at Radboud University and also include signs from the Chinese Sign Language corpus collected by Prof. Gong Qunhu at Fudan University in Shanghai. These data have been collected primarily through elicitation from two language informants. There is some lexical variation captured in the dataset; signs are predominantly listed with a basic English gloss, and in some cases an accompanying Chinese gloss.

> *Sign clips*

All sign clips were signed by fluent, deaf signers of NGT. All sign clips were between 2-4 seconds long, filmed on a neutral background. Signers did not use mouthing of spoken words in the recordings.[8] For the NGT signs, existing video clips from Signbank were used. For the CSL signs, new video clips were recorded

---

[8] In NGT, mouthings commonly accompany lexical signs (Bank et al., 2016), however given the comparative nature of the study and the focus on the manual component of the signs, mouthing were excluded.

based on existing clips in the database by a deaf signer of NGT. This was done to ensure uniformity in presentation of stimulus for the rating task, so that participants would produce ratings based solely on the form of the sign, and not be influenced by other factors such as signer ethnicity. While this means CSL signs were not recorded with native signers, the NGT signer who recorded these signs had some experience with Chinese Sign Language, having spent time in Shanghai collecting a portion of the Global Signbank CSL dataset and was able to fluently produce the CSL signs used as stimuli.

3.3. Phonological comparison

Using a method based the comparisons made by Yu et al. (2018) and Parks (2011), the lexical comparison combined a feature-based phonological analysis with a quantifiable measure of distance, creating a phonological overlap score for each cross-language sign pair. This score provided a basis for stimuli selection for the sign pairs used in the iconicity rating task.

Sign selection

Signs for the comparison were selected on the basis of a Swadesh list. Swadesh lists are concept lists developed for lexical comparisons that include core concepts that are present across cultures and thus form an appropriate dataset for cross-linguistic work. Originally developed for spoken languages, sign linguists have adapted Swadesh lists for work on signed languages (e.g. Bickford, 1991; Woodward, 1993a). The Swadesh list used for this study was developed for cross linguistic comparison by the ECHO project (Woll et al., 2010).[9]

Given the demands of the two-part design, including the time-consuming phonological coding of CSL signs required for the lexical comparison a subset of the 300-word Swadesh list items was selected. These were selected based on specific criteria of semantics, form and database availability. Firstly, the concepts selected for comparison were content words, roughly corresponding to nouns and verbs. Like in other lexical comparisons, I omitted concepts from the list corresponding to place names, numbers and pronouns (Ebling et al., 2015). Numbers and pronouns were not selected as they were likely to be produced across languages using either indexical pointing or manual counting. Country names were omitted as in a sign language, there is a tendency for these to be borrowed endonyms from the given country's local language. Secondly, concepts were selected based on the form of their corresponding signs; I included only monosyllabic signs. In addition to this, concepts whose form in either language incorporated some sort of written language; e.g. incorporation of the manual alphabet or manual representation of Chinese characters, were omitted.[10] Thirdly, concepts were selected based on their availability in the database; only concepts with video entries available in Global Signbank for each language were included in the comparison. After filtering for these criteria, a random selection of 95 concepts from the remaining list was used for the phonological comparison.

---

[9] The Swadesh list can be found online at http://sign-lang.ruhosting.nl/echo/.

[10] While fingerspelling is one form of iconic motivation in signs, for the purposes of the present study, I opted to omit this class of signs in line with other ratings studies (Vinson et al., 2008). In this way, judgements of iconicity can be interpreted independent of signers' familiarity with written language systems.

To identify pairs of translation equivalents, I performed a search in Global Signbank using the English translation of each Swadesh list concept to fist select the CSL sign. In many cases, this returned several CSL signs, some of which were lexical variants where two signs existed for the same concept. In these cases, one CSL variant was picked at random to represent the foreign sign for a given concept. In other circumstances, single concepts from the Swadesh list were split into two semantically distinct signs in CSL. An example of this is the concept *know*, which was listed in Signbank with two Chinese variants, 知道 meaning roughly *to be aware of*, and 认识 meaning roughly *to understand* or *to meet a person*. To select among the two, I examined the available signs in NGT with their translation equivalents, and with the help of a native Chinese speaker, selected the CSL-NGT sign pair that most closely aligned in meaning.

Next, a search was performed for the NGT translation equivalents. Where multiple NGT variants existed for a given Swadesh list concept, each variant was compared to the CSL sign. For example, given the concept *cat,* each of the four NGT variants found in Signbank for NGT-CAT were compared to CSL-CAT. Preserving such variation in the lexicon is important for a lexical comparison, as signers of a given language are likely to be familiar with different lexical variants even if they do not produce them (Johnston, 2003). The resulting set of CSL signs and their corresponding NGT signs formed the basis for the phonological comparison.

Phonological coding

Each of the signs selected for comparison was coded according to the detailed phonological coding scheme of Global Signbank (Crasborn et al., 2018b). For each sign, I used 20 of the phonological features encoded in Signbank for the comparison, only omitting fields that dealt with non-manual phonology (e.g. actions of the mouth) and factors related to corpus variation (e.g. weak-drop). A full list of the features and details of the coding scheme can be found below. The list of features were selected to provide a more comprehensive description of signs than provided by the formational parameters (see Yu et al., 2018) and to form the basis of a granular and quantifiable overlap score between signs (see Parks, 2011).

For each feature, a sign received either received one of the fixed values listed in the table below or was coded as not applicable (NA). Signs were coded as NA for a given feature if the particular feature was not necessary to describe the sign under the coding scheme. For example, a one-handed sign would receive an NA value for *Weak Hand Finger selection*, *Weak Hand Finger configuration*, *Weak Hand Spreading* and *Weak Hand Aperture*, as the weak hand is not necessary to describe the sign form. For certain fields, such as *Location* and *Handedness*, no signs were coded with NA values.

Table 1. Summary of the values used in the phonological coding scheme as outlined by Global Signbank manual (Crasborn et al., 2018)

Feature	Fixed Values
Finger selection[a][b]	I (Index), M (Middle), R (Ring), P (Pinkie), T (Thumb), IM, IMR, IMRP, TI, TIM, TIMRP, NA
Finger configuration[a]	Extended, Curved, Closed, Bent, NA
Spreading[a]	Spread, Unspread, NA
Aperture[a]	Open, Closed, NA

Handshape change	Wiggling, Spreading, Unspreading, Bending, Curving, Opening, Closing, NA
Location[b]	Chin, Shoulder, Weak Hand: back, Neutral space, Temple, etc.
Movement direction	Ipsilateral, Contralateral, Forwards, Backwards, NA
Movement shape	Arc, Circle, Zigzag, Spiral, NA
Repeated movement	True, NA (False)
Alternating movement	True, NA (False)
Relative orientation: Movement	Back, Base, Front, Finger tips, Radial, Ulnar, NA
Relative orientation: Location	Back, Base, Front, Finger tips, Radial, Ulnar, NA
Orientation change	Flexion, Extension, Rotation, Pronation, Supination, NA
Handedness	1 handed, 2 handed symmetrical, 2 handed asymmetrical, 2 handed non-symmetrical
Relation between articulators	Front, Back, Above, Below, Cross, Interlocked, Inside, NA
Contact type	Initial, Final, Double, Continuous, Brush, NA

[a] These features were coded for both the strong and weak hand in two handed signs.
[b] For the full list of fixed values for these fields, see the coding manual at https://signbank.science.ru.nl//about/manual/

Reliability

The current methodology provides the advantage of removing the subjectivity involved in judging similarity across signs, instead form-related decisions are made as a precursor to comparison (Yu, et al., 2018). In the initial phonological coding stage, the Global Signbank coding scheme attempts to reduce redundancy and standardise phonological descriptions by providing specific examples for feature values in the accompanying manual (Crasborn et al., 2018b). However, despite the comprehensiveness of the manual, several difficult cases were encountered in coding the CSL signs. Many of these were localised to specific fields, where either manual descriptions were vague or NGT examples in the database showed inconsistent coding. This was usually the case where the coding scheme was aimed to reduce redundancy, however the actual coded values encoded extra information. An example of this was the field *Relation between articulators* where many two-handed signs were coded as *front/back*, indicating the position of both articulators, when the coding only needed to reflect the dominant hand's position with relation to the non-dominant hand, thus a simple *front* or *back* would suffice. Given these challenges, while applying the Global Signbank coding scheme to the CSL signs, coding was reviewed with the help of an experienced coder. Review of the coding for NGT signs used in this study also revealed small inconsistencies in the database and these were corrected before the comparison was performed. Once all signs were satisfactorily coded for both datasets, the pairwise comparison was performed.

Procedure: Phonological comparison

To perform the phonological comparison, each sign was compared on the basis of the 20 coded features. In order to produce an overlap score for each pairwise comparison, the phonological comparison employed the rationale used in Levenshtein distance calculations. Instead of a measure of distance between two forms, the

comparison returned a measure of overlap for each sign pair. To do so, I used an automated algorithm to compare signs that carried out the following steps.

For each sign pair, each feature was examined. If two signs were coded with matching values, they received for this feature a similarity score of 1. If the two signs were coded with mismatching values, they received for this feature a similarity score of 0. If one sign was coded with a value and the other was coded with NA, they received a similarity score of 0. If both signs were coded with NA, the feature was dropped from the comparison. Thus, signs were only compared based on features that were specified with values for each sign. This was done so as to not inflate similarity scores due to the absence of a feature (and the resulting matching NAs). For example, any pair of one handed signs would align with NAs in all the *Weak Hand Handshape* features, thus appearing much more similar than any pair of two handed signs. Phonological overlap scores were normalised by dividing the similarity score by the number of features compared for each pair. Following Yu et al. (2018) all features were given equal weight in the comparison. The table below shows the procedure carried out in each pairwise comparison with the example of the signs CSL-APPLE and NGT-APPLE (Figure 7).
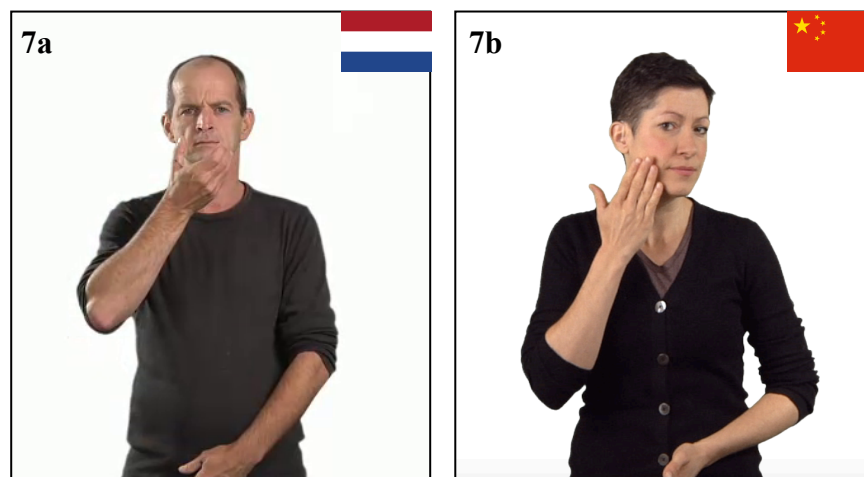

Figure 7. The signs NGT-APPLE (7a) and CSL-APPLE (7b)

Table 2. Calculation of phonological overlap score for CSL-apple and NGT-apple

Feature	APPLE CSL	APPLE NGT	Comparison	Similarity Pts.
Strong Hand Finger selection	TIMRP	IMRP	✗	0
Strong Hand Finger configuration	Extended	Curved	✗	0
Strong Hand Spreading	Unspread	Unspread	✓	1
Strong Hand Aperture	NA	Open	✗	0
Weak Hand Finger selection	NA	NA	-	-
Weak Hand Finger configuration	NA	NA	-	-
Weak Hand Spreading	NA	NA	-	-
Weak Hand Aperture	NA	NA	-	-

Handshape Change	NA	NA	-	-
Location	Cheek	Mouth	✗	0
Movement Direction	NA	Upwards	✗	0
Movement Shape	Circle	NA	✗	0
Repeated Movement	1	NA	✗	0
Alternating Movement	NA	NA	-	-
Relative Orientation: Movement	NA	Radial	✗	0
Relative Orientation: Location	Palm	NA	✗	0
Orientation Change	NA	Supination	✗	0
Handedness	1	1	✓	1
Relation between articulators	NA	NA	-	-
Contact Type	Continuous	NA	✗	0

Normalised overlap score: 2/13 = **0.154**

For each CSL sign, a pairwise comparison was performed with each NGT variant for the given concept. Thus, CSL-APPLE was compared to NGT-APPLE-A, NGT-APPLE-B, and so on. The pairwise comparison returned an overlap value for the CSL sign and each variant in the NGT lexicon; through this score it was possible to select the NGT variant with greatest phonological overlap with the CSL sign, which was used for the subsequent iconicity rating task.

3.4. Iconicity rating

In a task adapted from the studies of Adam et al., (2007) and Occhino et al. (2017), NGT signers were asked to rate signs from their native sign language, NGT, and signs from a foreign language, CSL. Sign pairs were matched by concept; for each concept participants viewed one NGT sign and one CSL sign.

Sign selection

In order to operationalise phonological overlap, the pairwise lexical comparisons were carried out, as detailed in Section 3.3. These comparisons provided a discrete overlap score for each pair of signs across languages. This phonological comparison also served the additional purpose of aiding in sign selection for the iconicity rating by selecting the most overlapping variant for a given CSL sign. In the case that a CSL sign had multiple NGT translation equivalents in Signbank, the NGT variant with the highest degree of phonological overlap would be selected for the rating task. From the 95 sign pairs used in the phonological comparison, 45 were selected for the rating task[11]. These pairs were selected based on their overlap scores, so that signs pairs with overlap scores ranging from 0-1 were included with a relatively even distribution in the stimulus set (mean = 0.48, SD = 0.3).

Participants

---

[11] See the appendix for the full list of signs and their Phonological Overlap scores.

Participants for this study were 11 adult signers of NGT (5 female), between 26 and 52 years old (mean = 34.2, *SD* = 8.9). All participants were living in the Netherlands and use NGT as their primary language of communication. Participants were recruited online through personal networks and social media. For each participant, I also gathered additional language background and demographic information.

Procedure: Iconicity rating

Participants participated in an online rating task based on the design of Occhino et al. (2017). All instructions were given in written Dutch. Participants were presented with 45 pairs of native and foreign signs. Each sign pair made up a block, with first the NGT sign presented, followed by the CSL sign. The 45 blocks of sign pairs were presented in randomised order.

Each sign was accompanied by several questions. For the critical responses variables of iconicity and similarity, participants were asked to respond on a sliding horizontal scale with two poles of 0 - *not at all* and 100 - *very much.* For foreign items, participants were presented with the CSL sign asked for only Iconicity Rating: (*How similar is the form of this sign to its meaning?*). For native items, participants were presented with the NGT sign and asked for both an Iconicity Rating, and a Similarity Score (*How similar is the form of this sign to the form of the Chinese sign?*). The two signs were each presented on their own page, but participants were free to navigate back and forth between signs. Below is an example of the iconicity rating, in English.[12]
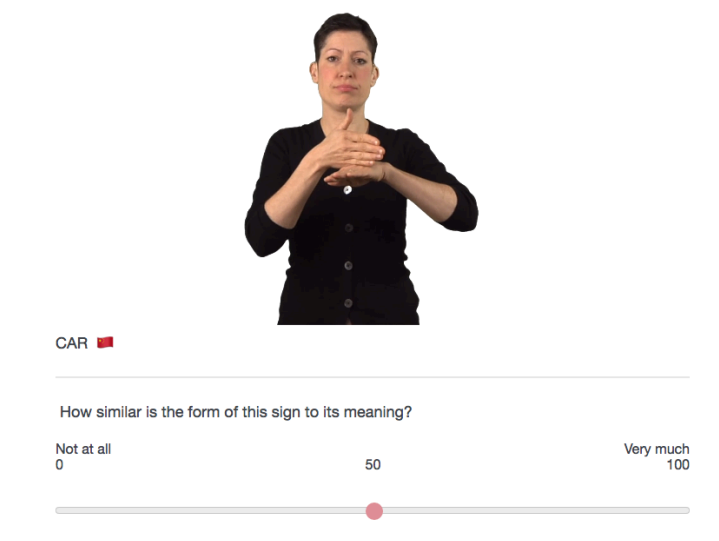

Figure 8. Example similarity rating (pictured in English)

For each sign, participants also answered an additional question concerning their usage and familiarity with the sign. In addition to these questions, participants filled out a background questionnaire to collect demographic data on age, language fluency and gender, as well as information on contact with other sign languages, and specific familiarity with Chinese/East Asian sign languages and culture.

---

[12] See the appendix for Dutch and English versions of all questions asked in the task

The duration of the task was approximately 40 minutes, and the task was administered using the online survey platform, Qualtrics. Nine out of the ten total participants completed the survey remotely via the internet, one completed it in the lab. As with other lengthly ratings studies (e.g. Vinson et al., 2008), participants were allowed to stop the task and resume at their convenience, as long as they completed the full task within an allotted time frame, in this case, 7 days.

3.5. Analysis

Each research questions was explored using linear mixed effects modelling. I analysed the data using the lme-4 package (Bates et al., 2014) in R version 3.1.4. (R development core team, 2017). The advantage of such analyses is that they can control for the random effects of both participants and items within the total observed variance, thus allowing for a more refined analysis of the effect of specific independent variables (e.g. Baayen et al., 2008). In keeping with this technique, I include random intercepts for both participant and item across all analyses. In addition to the mixed models, I also include more traditional statistical techniques to give a full overview of the effects in question (t-test, Pearson's correlation test).

4. Results

Overall, a great deal of variation was observed in the iconicity ratings gathered for pairs of translation equivalents in the present study, as illustrated in Figure 9. Many signs, both native and foreign, elicited a wide range of ratings from different participants. Despite this variation, a preliminary inspection of the data reveals two clear observations.

First of all, looking across the list, the NGT sign for a given concept received iconicity ratings that were on average equal to or higher than that of their CSL sign translation equivalent, with few exceptions such as *chair*, *look* and *mouse*. In each of these exceptional cases, however, the range of variation for native signs also included relatively high iconicity ratings. Overall, this provides preliminary evidence that NGT signers generally judge their native sign for a given concept to be more iconic than the foreign sign.

Secondly, despite the variation observed, there is also a substantial amount of systematicity. Specifically, certain concepts appear to have corresponding signs that are judged to be highly iconic across languages, with high agreement among signers such as *drink*, *road* and *hearing-aid*. In comparing these ratings to the phonological overlap scores (see appendix) it is apparent that many of these concepts have high overlap across languages. For sign-pairs that did not have high overlap scores, ratings were generally more spread out across participants, with variation observed for both CSL and NGT sign ratings (e.g. *hate*).
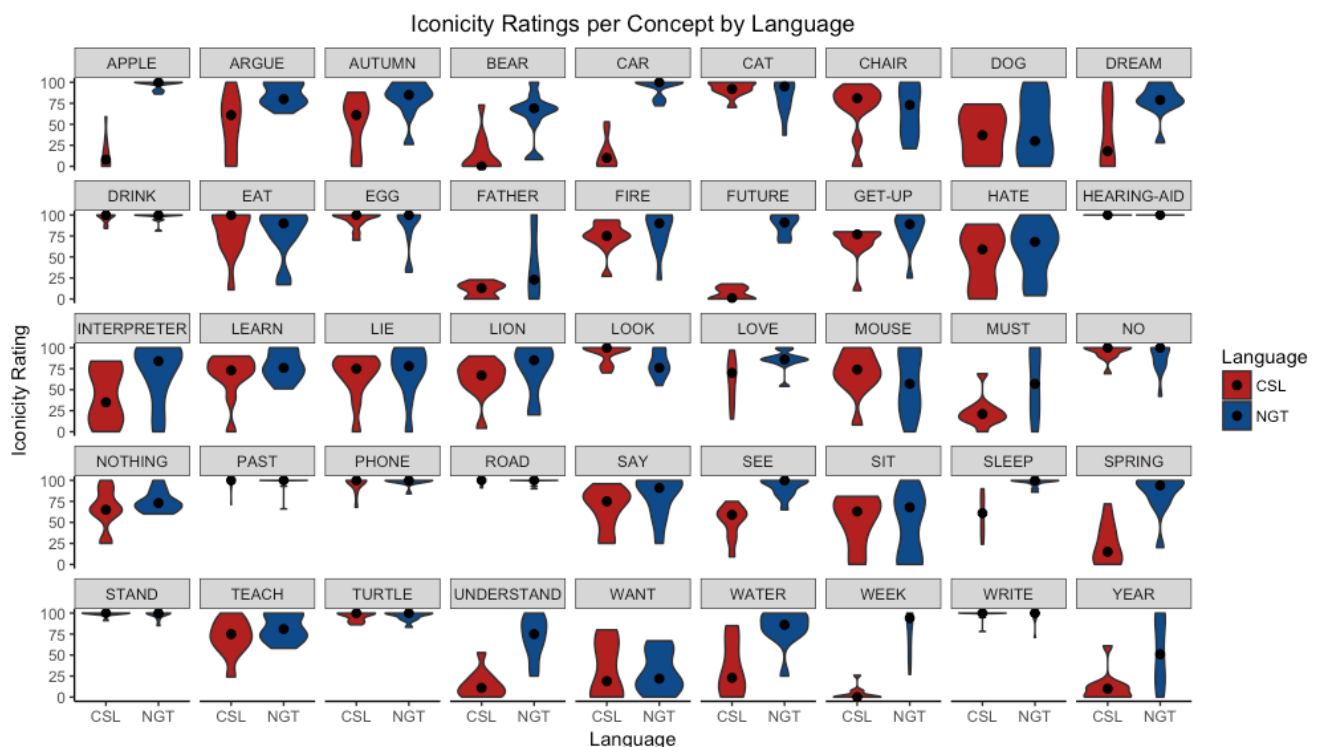

Figure 9. Iconicity ratings for translation equivalents across CSL and NGT

4.1. Usage

Given that I was interested in the effects of lexical knowledge on iconicity perception, it is critical that the sign variants selected were actually used by signers in the sample. The results from the question on usage

confirmed that the vast majority of NGT signs used in the rating task were familiar to signers in the sample. However, after reviewing participants' answers, I eliminated two sign pairs from the analysis based on their usage in NGT. The NGT signs selected for SIT and WANT in the stimuli were extremely low frequency variants among the signers in the sample, such that most signers reported never having seen these variants.

It is only useful to assess the contribution of lexical overlap on perceived iconicity when this overlap is with a variant that is actually in use; given that the effects of overlap are posited to stem from the repeated mapping of form to meaning in a given native sign. In fact, many NGT signers reported using/seeing CSL-WANT to mean ZIN IN or "fancy" (to fancy something). Again, given the effect under question, WANT was deemed inappropriate for this analysis. Thus, responses for these two sign pairs were removed from the data and the remaining 43 items were included in the analysis.

4.2. Foreign versus native iconicity

The first analysis assessed whether NGT signers judged native signs to be more iconic than their foreign translation equivalents. If native signs are judged to be more iconic than foreign signs, then NGT signers should produce higher iconicity ratings for NGT signs than CSL signs. A paired samples t-test compared the mean iconicity ratings produced by participants for each language. Results demonstrated that NGT signers rated CSL signs (m = 57.931, *SE* = 1.666) to be significantly less iconic than NGT signs (m = 78.341, *SE* = 1.260; *t* (10) = -4.868, *p* < 0.001). Thus, NGT signers show an effect of Language on mean iconicity rating.

Further analysis of the data using mixed models confirmed the effect of Language on Iconicity Rating. The Table 4 below summarises the best fit model to the dataset, which includes Language and Item as random effects and Language as a fixed factor. The inclusion of Language into the model significantly improved the explained variance from the null model (χ^2 = 14.73; *p* < 0.001). Thus, whether a sign came from CSL or NGT significantly predicted how iconic it was judged to be by participants.
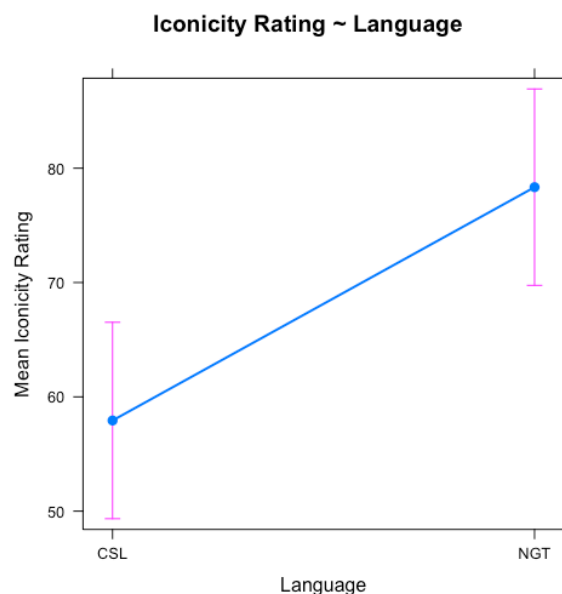

Figure 10: The effect of Language on mean Iconicity Rating found in Model 1

Table 3. The models used in the mixed model analysis of RQ1

Model	Formula
Null model	Iconicity Rating ~ 1 + (1|Participant) + (1|Item)
Model 1	Iconicity Rating ~ Language + (1|Participant) + (1|Item)

Table 4. Summary of the linear mixed effects model for RQ1, with Language as a fixed effect and Item and Participant (not included in the summary) as random effects.

Predictor	Estimate	*SE* (df)	*z*-value	*p*-value
(Intercept) (CSL)	57,931	4.378 (66.150)	13,234	<2e-16***
Language (NGT)	20,410	5.255 (88.00)	3,884	0.000199 ***

***$p < 0.001$, **$p < 0.01$, * $p < 0.05$

4.3. Lexical overlap and iconicity

The second analysis addressed whether the degree of lexical overlap between native and foreign signs could predict iconicity ratings of foreign signs. If iconicity ratings are influenced by a signer's lexical knowledge, then form-similarity with the native lexicon should influence iconicity ratings for foreign signs. Signs with higher overlap should receive higher iconicity ratings, and signs with lower lexical overlap should receive lower iconicity ratings. To examine this, I used two different measures of lexical overlap, a linguistic measure produced by the phonological comparison as well as a similarity rating produced by participants. Using linear mixed effects analysis, I examined the iconicity ratings produced by NGT signers for CSL signs.

A preliminary correlation analysis revealed a significant positive correlation between Phonological Overlap and Iconicity Rating ($r = 0.577$; p <0.001). Examining the data with mixed models further revealed Phonological Overlap to be a significant predictor of Iconicity Ratings in foreign signs. Including Participants and Items as random effects, I found the predictor of Phonological Overlap to significantly improve the null model. In examining the random effects of the null model and Model 2a, I found that the addition of Phonological Overlap as a fixed factor explained more than half of the variance observed across different Items. This suggests that more than half of the unexplained variation seen in different stimulus Items could be accounted for by their phonological overlap with their respective NGT sign. Examining the two models side by side, the new Model (2a) performed significantly better in explaining variance in the dataset than the null model ($\chi^2 = 30.176$, $p < 0.001$).

Similarity ratings produced by participants were also found to be positively correlated with Iconicity Ratings ($r = 0.737$, $p < 0.001$). Indeed, replacing the fixed effect of Phonological Overlap with the Similarity Rating produced an even stronger model (2b). The fixed factor of Similarity Rating accounted for a small portion of variance among Participants, however it increased the random variance in Items nearly threefold. This suggests that the variance in Iconicity Ratings produced by each participant can be explained in part by

examining the Similarity scores. As both are subjective ratings produced on the same scale, there may be observable systematicity in how individual participants produce the two ratings. The dramatic increase in variance per Item however, speaks to the wide range of unsystematic differences in Similarity Scores produced for each sign. Overall, Model 2b significantly improved the null model ($\chi^2 = 131.61$, $p < 0.001$) and was also significantly better than Model 2a in explaining variance in the dataset ($\chi^2 = 101.43$, $p < 0.001$).
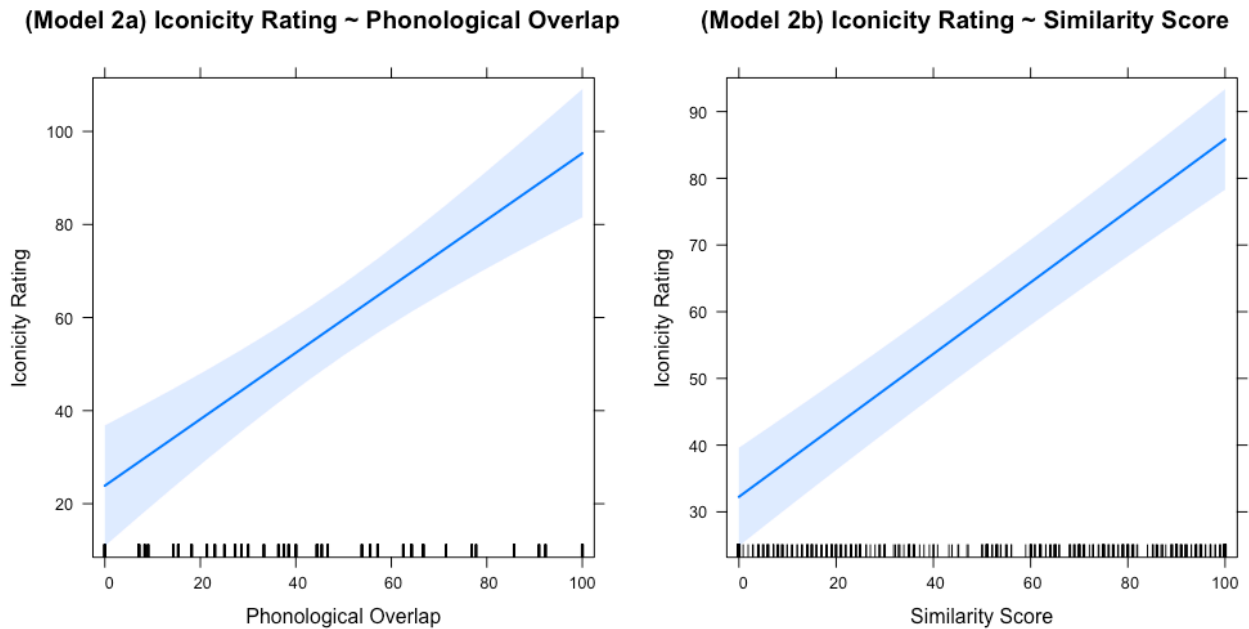

Figure 11: The fixed effects found in Models 2a and 2b

In an attempt to increase the explanatory power of the model I added both fixed factors of Similarity Score and Phonological Overlap to the model (2c). This revealed a correlation between the two variables ($r = 0.704$, $p < 0.001$).[13] Overall however, Model 2c demonstrated the best fit to the data, revealing that both Similarity Score and Phonological Overlap independently added significant explanatory power to the model.
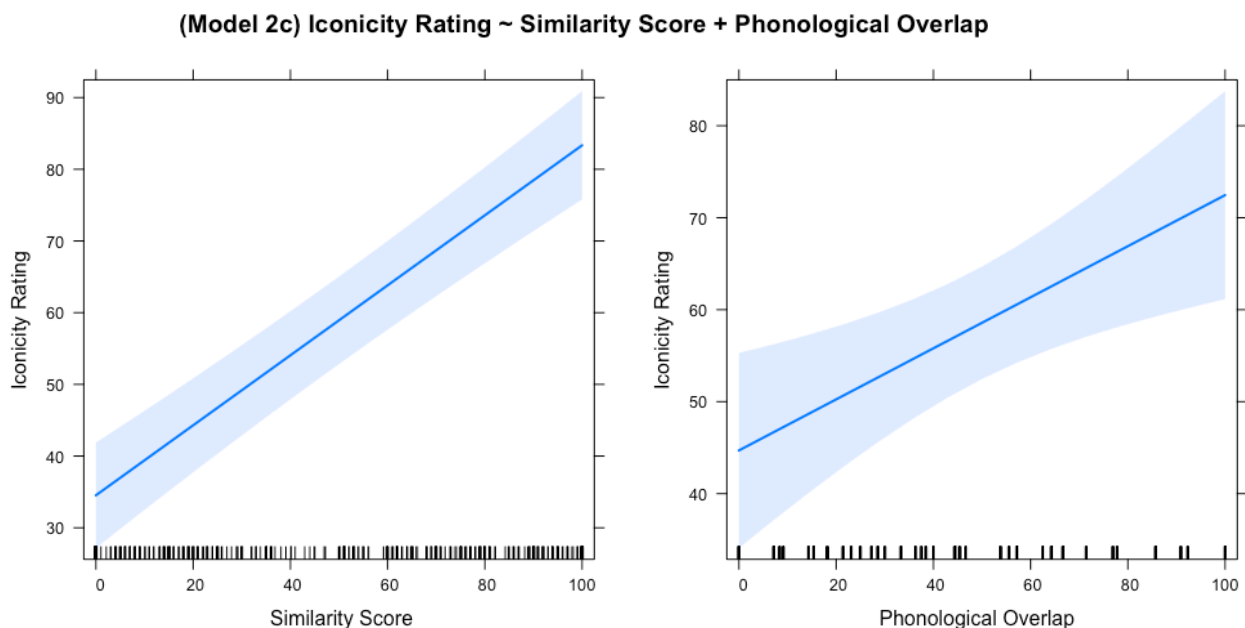
Figure 12: The fixed effects found in Model 2c

Table 5. The models used in the mixed model analysis of RQ2

Model	Formula
Null model	Iconicity Rating ~ 1 + (1|Participant) + (1|Item)
Model 2a	Iconicity Rating ~ Phonological Overlap + (1|Participant) + (1|Item)
Model 2b	Iconicity Rating ~ Similarity Rating + (1|Participant) + (1|Item)
Model 2c	Iconicity Rating ~ Similarity Score + Phonological Overlap + (1|Participant) + (1|Item)

Table 6. Summary of the linear mixed effects analysis for RQ2, with output from models 2a, 2b and 2c, including Similarity Score and Phonological Overlap as a fixed effects and Item and Participant (not included in the summary) as random effects.

Predictor	Estimate	*SE* (df)	*z*-value	*p*-value
Model 2a				
(Intercept)	23.8678	6.5839 (49.21)	3.625	0.000684 ***
Phonological Overlap	0.7143	0.1111 (43.00)	6.431	8.63e-08 ***
Model 2b				
(Intercept)	32,24660	3.74155 (59.70)	8,619	4.52e-12 ***
Similarity Score	0,53577	0.04016 (362.50)	13,341	< 2e-16***
Model 2c				
(Intercept)	21,29359	5.06857 (48.20)	4,201	0.000114***
Similarity Score	0,48816	0.04352 (485.80)	11,217	< 2e-16***
Phonological Overlap	0,27756	0.09258 (58.50)	2,998	0.003985**

***$p < 0.001$, **$p < 0.01$, * $p < 0.05$

5. Discussion

This study examined the effect of language experience, specifically lexical knowledge, in how signers perceive iconicity in foreign signs. Results demonstrated that NGT signers rate foreign CSL signs as less iconic than native signs. These results replicate the findings of Occhino and colleagues (2017) with similar effect sizes in two distinct sign languages, suggesting that signers show a robust native language iconicity bias when faced with foreign signs. Previous results from Occhino et al.'s (2017) two-way rating task demonstrate that this bias persists even in sign pairs where the foreign signs is rated as highly iconic by foreign language users. This, along with the considerable variation observed in the rating data collected in this study, remind us that it is difficult to objectively estimate any degree of "ground truth" iconicity. Additionally, given the random sampling method from a core vocabulary list, it seems unlikely that any alternative explanation may be driving the observed effect, such as a skewed selection that favours NGT signs with more visual mappings than their CSL translation equivalents. Thus, in line with the discussion of Occhino et al. (2017), I take these results to support the notion that iconicity is constructed within a network of language-internal patterns.

To explore the nature of these patterns, I further examined the role of lexical knowledge in perceived iconicity. Using two measures of lexical similarity, I found that indeed, the form-overlap between a given NGT sign and its CSL translation equivalent predicted iconicity. Using a linguistic measure of overlap based on the comparison of 20 phonological features, I found that CSL signs with a high phonological overlap score with their NGT translation equivalents were rated to be highly iconic. Interestingly, an even stronger predictor of CSL sign iconicity ratings was a measure of perceived similarity from signers. Using subjective ratings of form-similarity from participants, I found that CSL signs that were rated as looking highly similar to their NGT translation equivalent, were also rated to be highly iconic. Thus, both linguistic and perceived measures of lexical similarity predicted iconicity ratings of foreign signs.

In fact these two measures of phonological overlap and perceived similarity were correlated, indicating that this novel measure of distance between signs provided insight into how signers interpret sign similarity. This novel measure is useful as it relies on Levenshtein calculations to provide a discrete, quantifiable measure of overlap between two signs. Similar measures have been widely used in spoken languages, from predicting intelligibility across related languages (Beijering et al., 2008) to informing models of bilingual lexical processing (Dijkstra et al., 2010). In sign language research this measure promises equally broad applications, with possibilities to compare lexical overlap within languages, across languages, and even across modalities. This method improves on past lexical comparisons by returning a numeric score as opposed to a categorical distinction between *same*, *similar* and *different*. However, it is not without its limitations. Several challenges identified in this study must be dealt with to operationalise this method for cross-linguistic work.

Firstly, a recurrent problem of cross-linguistic comparisons must be faced, namely the issue of connecting concepts across languages. The field of lexicography has developed tools to minimise this challenge, with the development of standardised word lists such as Swadesh lists. However, even in the present study, using a Swadesh list specifically designed for sign languages, I encountered challenges where there was a one-to-

many correspondence of concepts across languages. Taking the example of the concept *know* (see Section 3.3), CSL and NGT carve up the semantic space differently, with NGT having a single concept and CSL (like spoken Chinese) splitting the concept into two distinct categories. Thus, to compare signs for know across languages, it is important to select concepts that are more or less mapped one-to-one across languages, or develop a method to select the appropriate corresponding concepts across languages. For the scale of the current study, it proved manageable to examine the available NGT sign(s) and its translation equivalents and select the Chinese concept that most closely formed a semantically aligned pairing with a given NGT sign. However, extending this method to comparing more languages at the same time, especially highly distinct languages, may introduce restrictions based on what lexical data is available for comparison, and how best to link concepts in an efficient, and possibly automated way.

Secondly, the present, the method allows for only the comparison of monosyllabic signs. However, in selecting signs, many Swadesh list concepts corresponded to compounds in CSL. This considerably narrowed the signs available for comparison. This is a challenge that can be fine tuned for future comparisons. For example, the algorithm can be adjusted to compare compounded and non-compounded signs much in the same way that it compares one and two-handed signs, by normalising the comparison across features compared (see Section 3.3, Table 2). Future studies may even choose to weight similarity between compounds and non-compounds in different ways, the purposes of comparison can inform changes to the algorithm.

Finally, the method relies on the presence of a lexical database that contains signs from different languages that are coded using the same phonological coding scheme. For the present study, this involved coding lexical signs from CSL according to the in depth phonological coding scheme outlined for the Global Signbank. This proved to be time consuming given the scale of the present, and imposed restrictions on how many signs were included in the lexical comparison. However, the presence of a unified lexical database that stores such information for various languages is an invaluable resource for cross-linguistic work. In this sense, the Global Signbank is unique, however it is also possible to also apply this overlap measure in comparing signs within a single language using language-specific lexical databases. In fact, the increasing development and use of lexical databases in sign language research can make this a useful tool for the future.

One additional challenge remained in using this phonological overlap score as a measure of lexical overlap between languages. To illustrate this, I will use the example of the sign CSL-EGG, an item with extremely high residual variation that was not explained by either measure of form-overlap.[14] For this sign, the measure of lexical overlap between NGT and CSL was low, however participants rated the CSL sign to be highly iconic. To understand the source of this variation, I examined the form of the CSL sign EGG, which depicts the cracking of an egg with two hands. This brings up two major sources of lexical knowledge that went unmeasured by the present method, but were revealed by examining participants' responses to the questions on usage. Firstly, several NGT signers reported using a sign of the same form as CSL-EGG to mean *omelette* in their native language. Thus, while the foreign sign may not directly overlap with the native translation equivalent, it overlapped with a concept that is closely related. In addition to this, one NGT signer reported

---

[14] See random effects residual plots in appendix.

using the CSL sign not as a lexical sign EGG, but to show the breaking of an egg (see Figure 13). This kind of usage can be best described as a classifier construction, construction in which the hands represent entities (in this case, human hands) depicting an action, movement or location. Classifiers constructions are not bound to the same rules, and constitute a more productive part of the sign language lexicon. However, as with the sign CSL-EGG, classifier constructions have the potential to become frozen to form fully lexicalised items (see Zwitserlood, 2008). In the case of this study, the contribution of classifier constructions to native lexical knowledge falls outside the measurable boundaries of language experience. My estimation of lexical knowledge is therefore restricted to the frozen lexicon as available in the Signbank database.
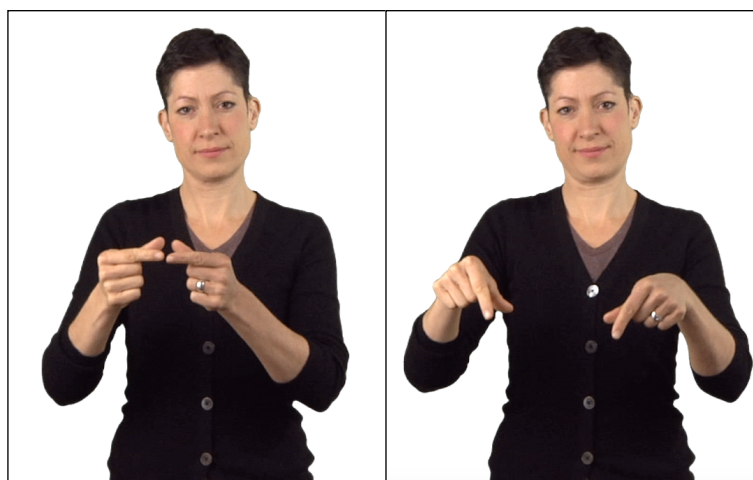

Figure 13. The sign CSL-EGG

Examining the residual variance left over from the model, it is clear that there are other signs whose variance cannot simply be explained by the factor of lexical overlap, or even the other factors that are related to language experience, such as the semantically related overlap or classifiers described above. Of course, as discussed in Section 2, various factors outside of language experience may play a role in how iconicity is perceived, such as culture, world experience. I do not attempt to offer lexical overlap as the sole predictor of iconicity, simply to put it forth as one of the factors that may contribute to constructing iconicity in a foreign sign.

Two potential issues in the presentation of the task were the use of an NGT signer to record CSL signs and the layout of the similarity judgement. The signing of CSL signs by a non-native signer introduced the possibility of NGT-accented production of CSL lexical items. Given that phonological differences between the two languages could lead to difficulty in production for NGT signer, the signer who recorded the signs was specifically chosen for her experience with CSL. This is however negligible as stimuli presented were isolated signs, and not signs in context. Furthermore, given the fact that this was a one-way rating task with no Chinese signers recruited for participation, this was not considered to be an issue for NGT signers. Another potential problem was that when producing similarity judgements, signers were presented with signs sequentially, on two separate pages, rather than simultaneously on the same page. While participants were able to navigate back and forth across translation equivalents to make this judgement, it is possible that this may have contributed to the large variation observed in similarity scores across participants. Future studies

could exploit adjusting the method to explore whether it decreases the observed variation. Despite this, the similarity scores produced by signers were highly correlated with the phonological overlap scores from the lexical comparison. Overall, neither of the aforementioned issues in presentation appeared to affect the results considerably, with similarity scores and iconicity ratings reflecting expected results.

Additional limitations of the current study include sample size and method of data collection. Initially, participants were recruited to perform the study online as a self-timed task over the period of 7 days. This design mirrored that of the BSL norming study performed by Vinson and colleagues (2008), where participants performed a lengthly rating task of 300 items over a period of 1 month. For the current task, while several participants elected to finish it all at once, other participants instead completed it over several days. In addition to this, one participant was recruited for participation in a lab setting, performing the entire task in one sitting, while taking short breaks in between. Along with these minor variations in data collection across the participants, the sample of NGT signers was also fairly small. Given that iconicity ratings can be subject to individual variation, a wider sample size would be ideal for examining the sources of the variation in depth. For example, a larger group would allow for the examination of the effects of age or multilingual exposure on iconicity ratings, while in the present sample these factors did not appear to be particularly informative. Nevertheless, the participants constituted a wide sample of the signing community, with a range of ages, varying regions and a balanced distribution of men and women. Despite these challenges in sampling and data collection the sample proved to be sufficient to capture the robust effects under investigation.

The results of the first two analyses provide good evidence for the effects of language specific experience on perceived iconicity. Unlike the bulk of research that focuses on evaluating differences between signers and non-signers in perceiving iconicity, this study focuses in on deaf sign language users, asking how their native-language experience can affect how they perceive iconicity in a foreign sign language. By examining signers, this study pinpoints a specific aspect of language experience, namely lexical knowledge, as a source of perceived iconicity. Like previous work, the results demonstrate a link between perceptual similarity of translation equivalents and iconicity ratings. This suggests that repeated use of the native lexicon reinforces specific iconic mappings between form and meaning; when signers recognise these mappings in foreign signs, they boost perceptions of iconicity. Furthermore, the correlation found between phonological overlap and iconicity ratings suggests that by quantifying these similarity relationships at a structural level, we can use linguistic measures of form-similarity to explain how signers perceive iconicity in a foreign sign. The results found in this study can be interpreted in line with usage-based approaches that connect our linguistic representations to factors such as familiarity and frequency. If iconic mappings are structures that connect semantic and phonological representations, then it follows that these linkages can be strengthened by repeated use. I propose that the relationship between phonological overlap and iconicity rating is a reflection of signers attending to familiar form–meaning associations foreign signs.

In fact, the connection between form-overlap and iconicity ratings does not come as a surprise. Previous results have demonstrated such a correlation; in both analyses of Adam et al. (2007) and Occhino et al. (2017), foreign signs that were judged to be highly similar to their translation equivalents were also judged to

be highly iconic. In fact, even in examining iconicity ratings produced by non-signers, Ortega and colleagues (2017) found non-signers rated NGT signs that overlapped with participants' gestural repertoire to be more iconic than signs that shared no overlap with gesture. By examining these effects with users of two sign languages, this study refines these observations by using what we know about the structural properties of signs to produce discrete measures of similarity by counting precise degree of formational overlap based on phonological criteria.

Furthermore, the suggestion that signers, and even non-signers interpretations of iconicity in foreign signs is mediated by overlap with their native lexicon, or even gestural repertoire can connect to findings from the spoken language realm. The processing of form-overlapping words found across languages, has been a rich source of research in the spoken languages, feeding into complex computational models of how non-native words are processed in the brain. Researchers have identified that higher form-similarity among cognates can lead to faster processing of these words (Dijkstra et al., 2010); and it has been proposed that native lexical knowledge, especially in low proficiency bilinguals, mediates access to foreign lexicon (Kroll & Stewart, 1994). So far, what we know about bilingual processing of these forms is limited to spoken languages. Furthermore, true cognates, forms that share etymological roots, are only present in languages that share history or contact, whereas false cognates, form-overlapping words across unrelated languages, are rare in spoken languages. Given that unrelated sign languages can share substantial form-overlap in their lexicon, there is much potential for novel investigation into how signers might process these cognates and false cognate forms from a foreign language. While acknowledging that signers faced with foreign lexical items are not strictly speaking low proficiency bilinguals, it is nevertheless possible that similar mechanisms may be at play, with the processing of foreign signs partially mediated by overlap with the native lexicon.

Predicting iconicity across languages by relying on linguistic properties of a sign can also elucidate the mechanisms of cross-language communication among deaf signers. While iconic signs have been identified as important in cross-signing settings to aid successful communication (Zeshan, 2015), however, what is considered iconic may differ across language communities. In this sense, simply labelling forms as *iconic* may not be the most insightful, without defining *iconic to whom*. Instead, it may be useful to examine the role of signs that recur across sign languages as formationally similar. This class of signs may contain recurrent iconic mappings tied to physical objects or actions, such as *write* in CSL and NGT, or may instead be less concretely tied to an image but still similar metaphorical mapping such as *past* in CSL and NGT. Above all however, these signs appear to be reliably perceived as iconic by signers of a foreign language. Thus, instead of the use of iconic signs as a factor in the success of cross-signing, a more concrete factor may be the use of lexicon that is form-overlapping across languages, as these are signs that are likely to be considered iconic by both parties. In fact, the possibility of linking lexical overlap to iconicity may be useful in predicting which forms are successful in cross-signing between signers with different language backgrounds.

Extending this to a broader lexical level, large-scale comparisons can create Levenshtein distances between sign languages, which may even be able to predict intelligibility or communicative success among signers of specific language pairs. In spoken languages, these measures have successfully correlated with mutual

intelligibility across related language pairs (e.g. Beijering et al., 2008), however, sign languages offer the possibility of examining intelligibility across totally unrelated languages. Parks (2011) performs an initial analysis of this using his Levenshtein calculations and data from intelligibility testing in the language varieties in his sample. He finds promising preliminary results, demonstrating a positive correlation between Levenshtein distance between language and performance on a one-way intelligibility task. In this way, we can apply measures of lexicostatistics to inform real-world settings, to see what linguistic distance really means at the level of communication.

These results can also inform experimental investigation into the effects of iconicity. Like previous work, they underscore that perceptions of iconicity can differ across groups of language users. This has implications for studies investigating how iconicity affects language processing in signers by operationalising iconicity ratings from groups of non-signers. If repeated use of form–meaning mappings strengthen iconic representations, then using non-signers as an estimate of objective iconicity may cloud important contributions of how features such as frequency, familiarity and usage can contribute to perceived iconicity. It is entirely possible that recurrence of iconic mappings within the lexicon may explain to some signs being perceived as more iconic than others; one example of this may be the mapping of the V handshape to eye gaze in NGT, in the signs LOOK, READ, SEARCH and SEE, as mentioned in Section 2.1. In addition to this, future work may also explore the role of familiarity and usage in perceived iconicity. Low frequency variants may be judged as less iconic than high frequency variants because repeated use can strengthen the links between form and meaning. Similarly, whether a signer herself uses a particular variant productively, or whether they are simply familiar with perceiving the variant may also affect iconicity ratings. Indeed, both within-language regularities and patterns of usage may boost iconicity ratings for signers, yet these same patterns may be lost on non-signers who lack the language experience to fortify these mappings. While this study looks across sign language lexicons, future work may look within lexicons to examine whether lexical knowledge can affect how signers perceive iconicity in their own native language.

6. Conclusion

Iconicity is a complex notion. On one hand, we can examine and trace the presence of specific mappings between form and meaning in signs. In fact we can even trace similarities in how these mappings recur across sign language lexicons. On the other hand, it cannot be denied that the perceived experience of iconicity is subject to individual biases. Cognitive linguists have attempted to bridge these two observations by pointing out that, like most aspects of language, the mappings between form and meaning that are involved in perceiving iconicity are mediated by the observer. This can result in differences in how iconic signs are judged to be, both the individual and community level, as experiences of the observer can affect the mapping process.

At the community level, one's native language can bias how iconicity is perceived. Studies have shown signers judge native signs as more iconic than foreign. This effect is mediated by lexical similarity, with foreign signs that resemble native translation equivalents judged to be more iconic than foreign signs that are formationally different than native signs. This study presents evidence that the degree of form overlap between a native an foreign sign, as measured by number of shared phonological features, can predict how iconic a signer judges a foreign sign to be. This relationship between phonological overlap and perceived iconicity can be interpreted as preliminary evidence that signers recognise familiar form–meaning mappings in a foreign sign. These familiar mappings are primed and strengthened by repeated use in the native language, thus seeing them in a foreign sign renders this sign more iconic.

This provides a first insight into how signers perceive foreign lexical items, a realm of investigation that will prove important as the work on cross-language communication among deaf signers steadily grows. It also adds to the building evidence that shows despite our intuitive ideas about what makes a sign iconic or not, there is much still to learn about how we construct these mapping. Nevertheless, through examining language experience, specifically lexical knowledge we can identify one source of systematicity in how iconicity is perceived.

References

Adam, M., Iversen, W., Wilkinson, E., & Morford, I. P. (2007). Meaning on the one and on the other hand. In C. Ljungberg, & O. Fischer (eds.), Insistent Images, 5, 211-227. Amsterdam: John Benjamins.

Al-Fityani, K., & Padden, C. (2010). Sign languages in the Arab world. In D. Brentari (Ed.), *Sign languages: A Cambridge language survey,* 433-450. New York, NY: Cambridge University Press.

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.

Bank, R., Crasborn, O., & Van Hout, R. (2016). The prominence of spoken language elements in a sign language. *Linguistics*, 54(6), 1281-1305.

Bates, D., Maechler, M., Bolker, B., Walker, S., Christensen, R. H. B., Singmann, H., & Grothendieck, G. (2014). Package 'lme4'. R foundation for statistical computing, Vienna, 12.

Battison, R. 1978. *Lexical Borrowing in American Sign Language*. Silver Spring, MD: Linstok.

Bradford, A., Sagara, K., & Zeshan, U. (2013). *Multilingual and multimodal aspects of "cross−signing" – A study of emerging communication in the domain of numerals*. Poster presented at TISLR11, London, UK.

Beijering, K., Gooskens, C., & Heeringa, W. (2008). Predicting intelligibility and perceived linguistic distance by means of the Levenshtein algorithm. *Linguistics in the Netherlands*, 15, 13-24.

Bickford, J. A. (1991). Lexical variation in Mexican sign language. *Sign Language Studies*, 72(1), 241-276.

Boyes-Braem, P. (1981). *Features of the hand shape in ASL*. Unpublished doctoral dissertation. University of California, Berkeley.

Brennan, M. (1990). Productive morphology in British Sign Language. *In Proceedings of the International Congress on Sign Language Research and Application, Hamburg*, 90, 205–228.

Brentari, D. (1998). A prosodic model of sign language phonology. Cambridge, MA: MIT Press.

Costa, A., Caramazza, A., & Sebastian-Galles, N. (2000). The cognate facilitation effect: implications for models of lexical access. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 26(5), 1283–1296.

Crasborn, O. A., van der Hulst, H., & van der Kooij, E. (2001). SignPhon: a phonological database for sign languages. *Sign Language & Linguistics*, 4(1), 215-228.

Crasborn, O. A., & van der Kooij, E. (1997). Relative Orientation in Sign Language Phonology. *Linguistics in the Netherlands*, 14(1), 37-48.

Crasborn, O., Bank, R., Zwitserlood, I., van der Kooij, E., Ormel, E., Ros, J., Schüller, A., de Meijer, A., van Zuilen, M., Nauta, Y. E., van Winsum, F. & Vonk, M. (2018a) NGT Signbank.10.13140/RG. 2.1.2839.1446.

Crasborn, O., Zwitserlood, I., van der Kooij, E., & Schüller, A. (2018b). Global SignBank manual. 10.13140/ RG.2.2.36603.82729.

Currie, A. M., Meier, R., & Walters, K. (2002). A crosslinguistic examination of the lexicons of four signed languages. In Meier, R., Cormier, K. & Quinto-Pozos, D., *Modality and structure in signed and spoken languages*, 224-236. Cambridge University Press.

Dijkstra, T., Miwa, K., Brummelhuis, B., Sappelli, M., & Baayen, H. (2010). How cross-language similarity and task demands affect cognate recognition. *Journal of Memory and Language*, 62(3), 284-301.

Dingemanse, M., Blasi, D. E., Lupyan, G., Christiansen, M. H., & Monaghan, P. (2015). Arbitrariness, iconicity, and systematicity in language. *Trends in Cognitive Sciences*, 19(10), 603-615.

Eccarius, P., & Brentari, D. (2010). A formal analysis of phonological contrast and iconicity in sign language handshapes. *Sign Language & Linguistics*, 13(2), 156-181.

Emmorey, K. (2014). Iconicity as structure mapping. *Philosophical Transactions of the Royal Society B*, 369(1651), 20130301.

European Sign Language Center (2018). Spread the Sign. Retrieved September 1, 2018 from https:// www.spreadthesign.com/

Gentner, D., & Markman, A. B. (1997). Structure mapping in analogy and similarity. *American Psychologist*, 52(1), 45-56.

Heeringa, W., Kleiweg, P., Gooskens, C., & Nerbonne, J. (2006). Evaluation of string distance algorithms for dialectology. In *Proceedings of the workshop on linguistic distances*, 51-62. Association for Computational Linguistics.

Hockett, C. F. (1960). The origin of speech. *Scientific American,* 203(3), 88-97.

Jarque, M. J. (2005). Double mapping in metaphorical expressions of thought and communication in Catalan Sign Language (LSC). *Sign Language Studies*, 5(3), 292-316.

Johnston, Trevor. (2003). BSL, Auslan and NZSL: Three signed languages or one? In Baker, A., van den Bogaerde, B., & Crasborn, O. (eds.), *Cross-linguistic perspectives in sign language research: Selected papers from TISLR 2000,* 47–69. Hamburg: Signum Verlag.

Klima, E. S., & Bellugi, U. (1979). *The signs of language.* Cambridge, MA: Harvard University Press.

KOMVA. (1989). *Basis-gebarenschat: een beknopt gebarenwoordenboek* [Basic sign language vocabulary: a concise sign dictionary]. Amsterdam: NSDSK.

Kroll, J. F., & Stewart, E. (1994). Category interference in translation and picture naming: Evidence for asymmetric connections between bilingual memory representations. *Journal of Memory and Language*, 33(2), 149-174.

Lepic, R., Börstell, C., Belsitzman, G., & Sandler, W. (2016). Taking meaning in hand. *Sign Language & Linguistics*, 19(1), 37-81.

Lerengebaren.nl. Retrieved September 1, 2018 from https://www.lerengebaren.nl/

Mandel M. 1977. Iconic devices in American sign language. In *On the Other Hand: New Perspectives on American Sign Language*, 57–108. New York: Academic

McKee, D., & Kennedy, G. (2000). Lexical comparison of signs from American, Australian, British and New Zealand sign languages. In K. D. Emmorey & H. Lane (eds.),*The signs of language revisited: An anthology to honor Ursula Bellugi and Edward Klima*, 49-76. Mahwah, NJ: Lawrence Erlbaum Associates

Meir, I. (2010). Iconicity and metaphor: Constraints on metaphorical extension of iconic forms. *Language*, 86(4), 865-896.

Occhino, C., Anible, B., Wilkinson, E., & Morford, J. P. (2017). Iconicity is in the eye of the beholder. *Gesture*, 16(1), 100-126.

Ortega, G., Schiefner, A., & Ozyurek, A. (2017). Speakers' gestures predict the meaning and perception of iconicity in signs. In G. Gunzelmann, A. Howe, & T. Tenbrink (Eds.), *Proceedings of the 39th Annual Conference of the Cognitive Science Society (CogSci 2017),* 889-894. Austin, TX: Cognitive Science Society.

Östling, R., Börstell, C., & Courtaux, S. (2018). Visual Iconicity Across Sign Languages: Large-Scale Automated Video Analysis of Iconic Articulators and Locations. *Frontiers in Psychology*, 9, 725.

Padden, C. A., Meir, I., Hwang, S. O., Lepic, R., Seegers, S., & Sampson, T. (2013). Patterned iconicity in sign language lexicons. *Gesture*, 13(3), 287-308.

Parks, J. (2011). Sign language word list comparisons: Toward a replicable coding and scoring methodology (M.A. Thesis). University of North Dakota, Grand Forks, ND.

Perniss, P., Thompson, R., & Vigliocco, G. (2010). Iconicity as a general property of language: evidence from spoken and signed languages. *Frontiers in Psychology*, 1, 227.

Pizzuto, E., & Volterra, V. (2000). Iconicity and transparency in sign languages: A cross-linguistic cross-cultural view. In K. Emmorey & H. Lane (Eds.), *The signs of language revisited: An anthology to honor Ursula Bellugi and Edward Klima,* 261-286. Manawah, NJ: Lawrence Erlbaum Associates.

R Core Team (2017). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

Sandler, W. (1989). *Phonological representation of the sign: Linearity and nonlinearity in American Sign Language*. Foris: Dordrecht.

Sandler, W. (2012). The phonological organization of sign languages. *Language and Linguistics Compass*, 6(3), 162-182.

De Saussure, F. (1916). *Course in General Linguistics*. New York: McGraw Hill.

Schermer, T., Harder, R., & Bos, H. (1988). *Handen uit de Mouwen: Gebaren uit de Nederlandse Gebarentaal in kaart gebracht* [Hands at work: Mapping signs from Sign Language of the Netherlands]. Amsterdam: NSDSK/Dovenraad.

Stokoe, W. C. (1960). Sign language structure. Studies in Linguistics Occasional Paper, 8.

Su, S. F., & Tai, J. H. (2009). Lexical comparison of signs from Taiwan, Chinese, Japanese, and American Sign Languages: Taking iconicity into account. In Tai, J. & Tsay, J. (Eds.), *Taiwan Sign Language and Beyond*, 149-176.

Taub, S. F. (2000). Iconicity in American sign language: concrete and metaphorical applications. *Spatial Cognition and Computation*, 2(1), 31-50.

Taub, S. F. (2001). *Language from the body: Iconicity and metaphor in American Sign Language*. Cambridge: Cambridge University Press.

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

Van der Kooij, E. (2002). *Phonological categories in Sign Language of the Netherlands. The Role of Phonetic Implementation and Iconicity.* LOT, Utrecht.

Vinson, D. P., Cormier, K., Denmark, T., Schembri, A., & Vigliocco, G. (2008). The British Sign Language (BSL) norms for age of acquisition, familiarity, and iconicity. *Behaviour Research Methods*, 40(4), 1079-1087.

Wilcox, S. (2004). Cognitive iconicity: Conceptual spaces, meaning, and gesture in signed language. *Cognitive Linguistics*, 15(2), 119-148.

Woll, B., Crasborn, O., van der Kooij, E., Mesch, J. & Bergman, B. (2010). Extended Swadesh List for Signed Languages. Online resource, ECHO Project. http://www.let.ru.nl/sign-lang/echo/

Woodward, J. 1993a. Lexical evidence for the existence of South Asian and East Asian sign language families. *Journal of Asian Pacific Communication,* 4.2, 91-106.

Yang, J. (2015) Chinese Sign Language. In Bakken Jepsen, J., de Clerck, G., Lutalo-Kiingi, A., & Mc.Gregor, W. (Eds.), *Sign languages of the World: A Comparative Handbook*, 177-194. Berlin: De Gruyter Mouton.

Yu, S., Geraci, C., & Abner, N. (2018). Sign Languages and the Online World Online Dictionaries & Lexicostatistics. I*n Proceedings of the Eleventh International Conference on Language Resources and Evaluation (LREC 2018)*, 4235-4240.

Zeshan, U. (2015). "Making meaning": Communication between sign language users without a shared language. *Cognitive Linguistics*, 26(2), 211-260.

Zwitserlood, I. (2008). Morphology below the level of the sign - frozen forms and classifier predicates. In J. Quer (Ed.), *Proceedings of the 8th Conference on Theoretical Issues in Sign Language Research (TISLR),* 251-272. Hamburg: Signum Verlag.

Appendix

Table 7. List of Swadesh list items used for iconicity rating task, with the ID-glosses for each sign as it is appears in Global Signbank and their phonological overlap score

Swadesh List Concept	CSL-gloss	NGT-gloss	Phonological Overlap Score
chair	CHAIR	STOEL	0
sit (v.)	SIT	ZITTEN-B	0
car	CAR	AUTO	0,007142857
tortoise	TORTOISE	SCHILDPAD-A	0,007142857
hate	HATE	HAAT	0,008333333
write	WRITE	SCHRIJVEN-B	0,009090909
get up	GET-UP	OPSTAAN	0,014285714
apple	APPLE	APPEL	0,015384615
must (v.)	MUST	MOETEN-A	0,015384615
dream (v.)	DREAM	VOORSTELLEN-A	0,018181818
egg	EGG	EI-A	0,021428571
love	LOVE	HOUDEN-VAN	0,023076923
water	WATER	WATER-A	0,025
road	ROAD	WEG-A	0,027272727
bear	BEAR-A	BEER-B	0,028571429
dog	DOG	HOND-B	0,03
fire (n.)	FIRE	BRAND	0,033333333
see	SEE-A	ZIEN-A	0,036363636
father	FATHER	VADER-A	0,0375
interpreter	INTERPRETER	TOLK	0,038461538
learn	LEARN	LEREN-A	0,04
future	FUTURE	TOEKOMST-A	0,044444444
want	WANT	WILLEN-B	0,044444444
eat	EAT-C	ETEN-B	0,045454545
argue	ARGUE	RUZIE-A	0,046666667
nothing	NOTHING-A	NIKS-C	0,053846154
lie (v., to tell a lie)	LIE	LIEGEN	0,055555556
autumn	AUTUMN	HERFST	0,057142857
understand	UNDERSTAND	BEGRIJPEN	0,057142857

Swadesh List Concept	CSL-gloss	NGT-gloss	Phonological Overlap Score
say	SAY	ZEGGEN	0,0625
mouse	MOUSE	MUIS-B	0,064285714
teach (educate)	TEACH	LESGEVEN	0,064285714
cat	CAT	POES-C	0,066666667
year	YEAR	JAAR-B	0,066666667
spring	SPRING	LENTE	0,071428571
lion	LION	LEEUW-A	0,076923077
past (time)	PAST	VROEGER-A	0,077777778
hearing aid	HEARING-AID	GEHOORAPPARAAT-B	0,085714286
look	LOOK	KIJKEN-A	0,085714286
sleep (v.)	SLEEP	SLAPEN-A	0,090909091
week	WEEK	WEEK	0,092307692
drink (v.)	DRINK	DRINKEN-A	0,1
no (say no)	NO	NEE-E	0,1
telephone	TELEPHONE	TELEFOON-A	0,1
stand (v., to be standing)	STAND	STAAN-C	0,1

Figure 14. Example of the rating task for both CSL (left) and NGT (right) signs LEARN

Original Dutch questions:

Iconicity: Hoe vergelijkbaar is de vorm van
dit gebaar met de betekenis ervan? (*How similar is the form of this sign to its meaning?*)

Similarity: Hoe vergelijkbaar is de vorm van dit gebaar met de Chinese gebaar? (*How similar is the form of this sign to the Chinese sign?*)

Usage: Gebruik je dit gebaar? (*Do you use this sign?*)

Ja, met dezelfde betekenis. (*Yes, with the same meaning.*)
Ja, met andere betekenis. Zo ja, leg uit. (*Yes, with a different meaning. If so, explain.*)
Nee, mar ik heb dit gebaar gezien met dezelfde betekenis. (*No, but I have seen this sign*

with the same meaning)

    Nee, en ik heb dit gebaar nooit nog gezien. (*No and I have never seen this sign*)

Table 8. The mixed models used in the analysis of the relationship between Similarity Score and Phonological Overlap

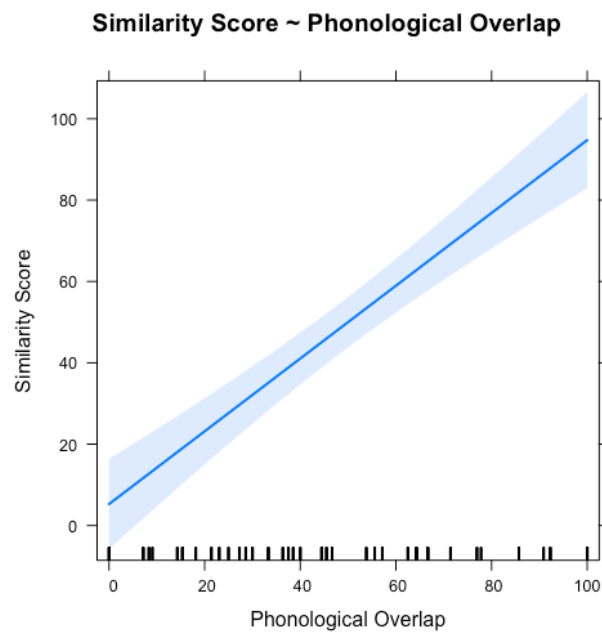| Model	Formula
Null model	Similarity Score ~ 1 + (1|Participant)
Model 3	Similarity Score ~ Phonological Overlap + (1|Participant)

Figure 15: Fixed effect of Phonological Overlap on Similarity Score as found in Model 3

Table 9. Summary of Model 3 with Phonological Overlap as a fixed effect and Item and Participant (not included in the summary) as random effects.

Predictor	Estimate	*SE* (df)	*z*-value	*p*-value
(Intercept)	2.33781	2.46628 (68.3)	948	347
Phonological Overlap	0.94872	0.04168 (419.0)	22.761	<2e-16***

***$p < 0.001$, **$p < 0.01$, * $p < 0.05$

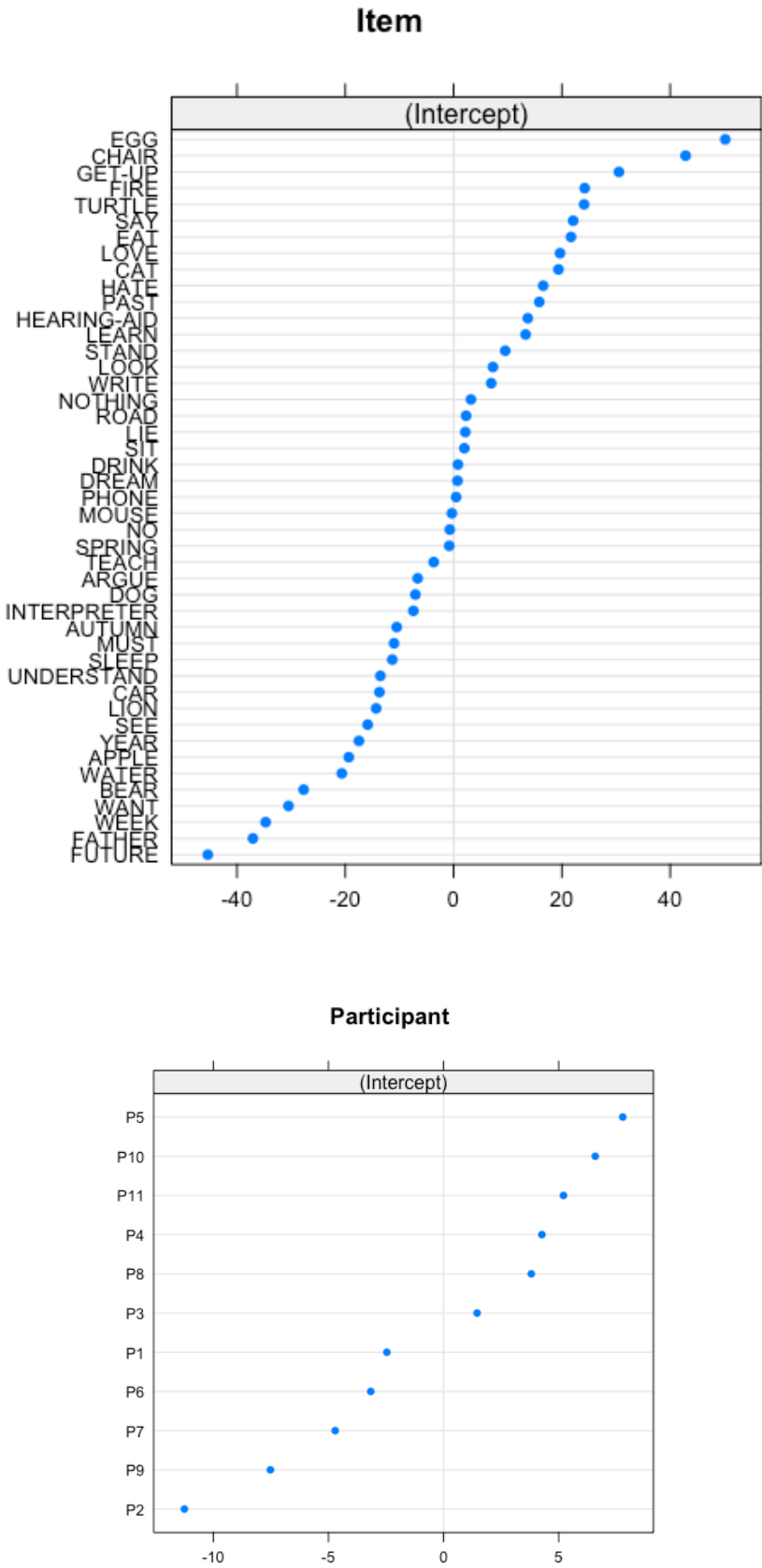Figure 16: Residual plots for random effects of Item and Participant in Model 2a

Figure 17: Residual plots for random effects of Item and Participant in Model 2b