The contribution of individual parameters to perceived iconicity and transparency in gesture-sign pairs

Door Spruijt (University of Cologne)
dspruijt@uni-koeln.de

Pamela Perniss (University of Cologne)
pperniss@uni-koeln.de

Petra B. Schumacher (University of Cologne)
petra.schumacher@uni-koeln.de

Abstract

It is often assumed that gestures are more iconic than signs, as they do not have to conform to a linguistic system. This study introduces an expanded methodology to explore (a) the relative transparency and iconicity of silent gestures and signs, and (b) the iconicity of three individual parameters (handshape, location and movement). We elicited meaning guesses and iconicity ratings (both whole-item and for each parameter) from sign-naive participants for both gestures and signs. Pilot data provide no evidence for differences in transparency and iconicity of gestures and signs, but we do find interesting examples of signs rated as more iconic than gestures. The iconicity of all three parameters is correlated with the iconicity of the whole item in both gestures and signs, but there may be a role for iconic strategies and the saliency of individual parameters. With this method, we provide a novel, more fine-grained manner of investigating iconicity in the manual modality.

Keywords: Iconicity, Transparency, Parameters, Sign language, Silent gesture

1 Introduction

In this paper, we present pilot results of a methodology that provides us with empirical evidence to test a prevalent assumption in the field of iconicity in the visuo-spatial modality: namely, that silent gestures (i.e. gestures produced by non-signers in the absence of speech) are highly iconic (e.g. Poggi 2008), and, in fact, more iconic than signs (implied by e.g. Frishberg 1975). Tkachman (2023) even explicitly states this as a hypothesis in her discussion of Ortega, Schiefner, and Özyürek's (2019) study on the perception of iconicity in (silent) gesture-sign pairs. This hypothesis seems plausible when considering that signs have to conform to the restrictions of a phonological system, likely abstracting the exact form of the referent at least somewhat, whereas gestures are not constrained in this way and are free to map referents onto the hands and body as directly as possible. The relative iconicity of gestures and signs has, however, never been quantified or empirically attested. By applying

Spruijt, Door, Pamela Perniss & Petra B. Schumacher (2023). The contribution of individual parameters to perceived iconicity and transparency in gesture-sign pairs. *FEAST* 5: 194-209.
https://doi.org/10.31009/FEAST.i5.16 http://www.raco.cat/index.php/FEAST

methods used to measure and operationalize iconicity in sign languages, we provide a way to quantify and directly compare the relative iconicity of gestures and signs.

In addition, we expand the existing methodology of iconicity ratings by applying it not only to whole (lexical) items, but also to sublexical units. This allows us to investigate the contribution of these units to the perceived iconicity of gestures and signs, providing deeper insight into the interplay between phonology and iconicity. The contribution of individual parameters to iconic representations has been regularly acknowledged in the literature, but any quantification of this has been mostly restricted to the intuitions of the researchers themselves (Cates et al. 2013; Pietrandrea 2002; Poggi 2008). By asking people to rate the iconicity of individual parameters, we move away from the biases of individuals, and towards empirically-driven quantification of iconicity in sublexical units.

2 Background

2.1 Iconicity and Transparency

Iconicity and transparency, while related, are distinct notions. Iconicity is the presence of a resemblance between a form and its meaning, whereas transparency is the accessibility of a meaning through form alone (e.g. Occhino et al. 2017; Perniss, Vigliocco, and Vinson 2014; Sevcikova Sehyr and Emmorey 2019). Iconicity is typically assessed via resemblance ratings, while transparency is reflected in the elicitation of meaning (see below for further details). So while highly transparent items are iconic in nature, as something in the form gives away the meaning, highly iconic items are not necessarily transparent. In our data (see Section 4.2), the DGS-sign PHOTOGRAPHY (Figure 1) is both iconic and transparent, while the sign COAT (Figure 2) is rather iconic, but non-transparent.

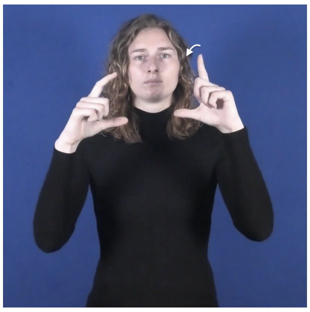

Figure 1: DGS-sign PHOTOGRAPHY.

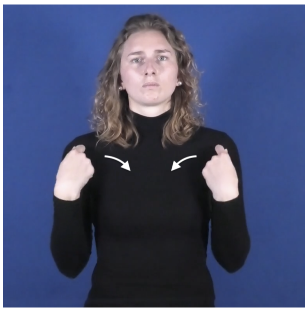

Figure 2: DGS-sign COAT.

The most frequent method used to determine the transparency of a lexical item is to ask people who are unfamiliar with the source language to guess the item's meaning. The proportion of participants who guess correctly is taken to reflect the transparency of the item. Recently, new measurements have been proposed by Sevcikova Sehyr and Emmorey (2019), who additionally ask for a rating of how obvious the guesser thinks the provided meaning would be to others ('perceived transparency'). Furthermore, they analysed not only the correct meaning guesses in a transparency score, but also looked at the diversity (or lack thereof) of guessed meanings for each item, regardless of correctness. These two measures combined give a measure of what they call the 'semantic potential' of an item, or how strongly it evokes a consistent meaning.

In early work on the iconicity of signs in sign languages, researchers tended to assign an iconicity 'status' to an item, usually resulting in the binary distinction of iconic and arbitrary items. The underlying assumption of this method, however, is that iconicity is binary, and a researcher can objectively determine to which class a sign belongs (Occhino et al. 2017). More recent insights, namely that iconicity is scalar, rather than binary, and is strongly mediated by the personal (linguistic) experience of an individual, gave rise to the use of ratings by a group of participants (see Dingemanse, Perlman, and Perniss 2020 for a review of different construals of iconicity). This method of determining iconicity also opened up the possibility to test the influence of human, cultural or linguistic experiences on the perception of iconicity. As such, the familiarity with a visuo-spatial language (Sevcikova Sehyr and Emmorey 2019, Trettenbrein et al. 2021) and knowledge of the specific sign language (Occhino et al. 2017; Omardeen 2018) have been shown to influence the perception of iconicity.

The perception of iconicity in an unfamiliar sign language has been moreover found to be influenced by the degree of form overlap with a sign or gesture the perceiver is familiar with. For signers, this would be a sign with the translation equivalent from a sign language they know (Omardeen 2018), and for hearing sign-naive people it would be a gesture expressing the same meaning from their gestural repertoire (Ortega, Schiefner, and Özyürek 2019; Ortega, Özyürek, and Peeters 2020). The availability of such a gestural repertoire has been assumed in previous research, with culturally-specific emblems (Poggi 2008, p.56; Ortega and Morgan 2015, p.455) or possible gestures someone might use to express a meaning (Goldin-Meadow and Brentari 2017, p.10) serving as a comparison to signs. However, these gestures are usually included for illustrative purposes, and not part of the research design itself. In recent years, studies have started to do just that: make the gestural repertoire explicit by eliciting silent gestures and then studying it (e.g. Ortega and Özyürek 2020; Van Nispen, Van De Sandt-Koenderman, and Krahmer 2017).

2.2 Silent gesture

Silent gestures, occasionally referred to as pantomimes(e.g. Van Nispen, Van De Sandt-Koenderman, and Krahmer 2017), used to study the gestural repertoire of non-signers, differ from the gestures used in everyday spoken language communication . Silent gestures, as their name implies, are produced in the absence of speech, and therefore carry the entire communicative load, whereas co-speech gestures support or enhance the communicative information encoded in speech (McNeill 2000). In this way, silent gestures are more like signs in a sign language than co-speech gestures are, and thus offer a better source of comparison for visual iconicity. The drawback is that silent gestures are not typical of natural communication, and as such have to be elicited in experiments to gather enough of them (Van Nispen, Van De Sandt-Koenderman, and Krahmer 2017).

Experiments aimed at eliciting these ’labels for referents’ entail the presentation of stimuli to hearing, sign-naive participants, who are asked to silently convey the intended meaning. The form of these stimuli can range from videos or illustrations to written words, depending on the purpose of the study. In general, sign language research tends to prefer the use of non-linguistic stimuli to minimize linguistic influence. However, in the investigation of iconicity, a written stimulus may be preferred to avoid the influence of the chosen visual representation on the form-meaning mapping. For example, imagine two pictures depicting a pig, one in which the nose is very prominent, and one in which the pig’s curled tail is salient. The first may lead to silent gestures related to the nose, while the second may make gesturers focus on the tail.

The iconicity or transparency of (silent) gestures is often assumed to be high, since they are treated as non-conventionalized, non-codified productions (e.g. Poggi 2008; Tkachman

2023). In cases where gestures do conventionalize, they form a new category: emblems. These are gestures that are culturally determined and shared across a community of users, and are relatively frequent in everyday use. They distinguish themselves from co-speech gestures by having a stable meaning that is understood with or without speech, rather than an ad hoc meaning in a spoken context. Emblems are clearly part of the gestural repertoire of a spoken language population, but recent elicitation of silent gestures suggests that there are other gestural representations that seem to be shared within a group, even without conventionalized meaning and frequent use. Ortega and colleagues (2019; 2020) show that there are many regularities in how silent gesturers express certain concepts. For more than a third of the concepts presented to participants, the majority of participants produced a very similar gesture (what Ortega et al. call a 'systematic gesture'). Furthermore, in many semantic domains, there is a preference for a certain iconic strategy (i.e. mapping strategy) across participants. Van Nispen, Van De Sandt-Koenderman, and Krahmer (2017) further show that silent gestures that have what they call a 'default' iconic strategy (i.e. more than 80% of the participants use a certain strategy), are more easily understood than those that do not have such a default strategy. Thus, while the freedom in producing gestures (without a linguistic system to conform to) has often been cited as a reason for gestures being very iconic, we can find more and more regularities in how concepts are referred to in silent gestures, indicating that perhaps gestures, like signs, vary in their iconicity.

To date, there is very little empirical evidence for the iconicity or transparency of gestures, let alone relative to signs. Van Nispen, Van De Sandt-Koenderman, and Krahmer (2017) measured the transparency (which they refer to as the comprehensibility) of gestures in both an open-answer and forced-choice task, and demonstrated that choice of method strongly affects the results; the forced-choice task showed ceiling effects while the open-answer task resulted in very few correct guesses. Ortega and Özyürek (2020) set the first step in the direction of quantifying the relative iconicity of gestures and signs, by showing that iconicity ratings[1] can be obtained for systematic gestures. Relevant to the current study is the fact that the iconicity of the gestures was rated using almost the entire scale, ranging from 1.78 ("pram/stroller") to 7.0 ("to wring" and "to clap") on a 7-point scale, indicating that gestures can also be relatively non-iconic. Such a low iconicity score for a gesture suggests that gestures may not necessarily (significantly) exceed that of signs.

2.3 Sublexical units

Gestures are generally regarded as holistic items that have no internal structure (Goldin-Meadow and Brentari 2017; McNeill 2000; Perniss, Vigliocco, and Vinson 2014). Signs on the other hand, have an internal structure; they are made up of smaller sublexical parts that combine to form meaning. The generally agreed upon main parameters of a sign are the hand configuration, location (or place of articulation) and movement. The orientation of a sign is sometimes seen as a fourth parameter (Battison 1978), yet is in other cases considered part of the handshape parameter (Sandler 2011). Recently, more and more attention is given to smaller sublexical units, the features (e.g. selected fingers, movement direction or repetition of movement), over the use of parameters as overarching units. This has the benefit that features can be described that do not neatly align with a single parameter, such as the

1. The instructions in their study perhaps do not strictly capture only iconicity, as Ortega and Özyürek (2020) asked participants to rate how well a gesture depicted the meaning, which narrows the scope to depictive resemblances, rather than the entirety of iconic relations. They further refer to the iconicity ratings as measures of "meaning transparency", which muddies the distinction between iconicity and transparency further.

relation between the two manual articulators, which pertains to both the handshape and the location.

The contribution of sublexical units to perceived iconicity has been sporadically investigated. Pietrandrea (2002) takes the iconicity of the parameters handshape and location as a measure of the incidence of iconicity in the lexicon of Italian Sign Language (LIS), showing that iconicity is not restricted to whole signs, but can be found in sublexical items as well. Poggi (2008) takes the argument further, by proposing that the more parameters in a gesture or sign are iconic, the greater the iconicity of the whole item is. As to the relative contribution of each parameter, Cates et al. (2013) found that location is the biggest contributor to iconicity in ASL (American Sign Language) signs. Importantly, these studies all coded both whole items and parameters as iconic or arbitrary, maintaining the traditional binary interpretation of iconicity.

Given the fact that iconicity can be found on the sublexical level, and contemporary best practice is to elicit iconicity ratings, we aim to gather more fine-grained information on sublexical iconicity by eliciting iconicity ratings for individual parameters. We chose to focus on parameter rather than feature ratings for two reasons. Firstly, we do not expect untrained participants to be able to recognize features reliably, both because they are small parts of a gesture or sign, and because there are far more features that they would have to remember the definition of than there are parameters. Secondly, as a first foray into the iconicity of sublexical units, we take the units that are compatible with the most models and as such, well-described and uncontroversial: handshape, location and movement.

3 Method

3.1 Gesture-sign pair selection

Gestures were elicited from 234 word prompts, covering a range of semantic categories as well as concrete and abstract concepts[2]. Sixteen German hearing, sign-naive participants (13 female) were shown the prompts and asked to produce a gesture that conveyed that meaning. They were told not to point to referents in the room, not to speak during gesture production, and were given the option to pass if they felt they could not produce a gesture for a concept. Each concept was shown for 4 seconds, during which participants would gesture their response. After each item a fixation cross (0.5 seconds) would reset their focus for the next item. The participants were filmed with a frontal and side camera.

As we are interested in the gestural repertoire of sign-naive speakers of German, we first selected only the gestures that had a production agreement of more than 50%: that is, more than half of the participants produced a gesture with (i) the same iconic motivation and (ii) a more or less stable form (Ortega and Özyürek 2020). As opposed to iconic strategies, which are generalized categories for the mapping between the form and referent (e.g. 'acting (on)', 'representation') that can describe any sign or gesture, iconic motivations are specific to reference to a particular concept, i.e. a single gesture or a sign. For example, in Figure 3 we see two separate gestures for the concept BASKETBALL, which share the iconic strategy 'acting (on)' because the hands represent the manipulation of an entity. However, how these gestures are motivated, i.e. which iconic image they map onto, differs. The left image shows the enactment of bouncing the basketball on the ground, whereas in the right image the gesturer maps throwing the ball (presumably towards the basket). Here, we looked for gestures

2. The gestures were originally elicited in a different project (Schiefner, Perniss, and Ortega 2023, forthcoming).

that shared an iconic motivation, which highly likely means they also share a strategy. The form of these gestures, from a sign language (i.e. phonological) perspective, can vary quite substantially (e.g. selected fingers extended or curved). For a gesture's form to be considered stable, we required it to have multiple overlapping features in each of the three main parameters (handshape, location and movement), but with leeway similar to the range of allophonic variation in sign languages. When both conditions (shared motivation and stable form) were met, the gestures were included in the set of "predominant gestures". 107 out of the 234 elicited concepts yielded such a predominant gesture.

Figure 3: Two gestures for BASKETBALL, sharing the iconic strategy 'acting (on)' but with different iconic motivations: bouncing the ball (left), and throwing the ball (right).

These predominant gestures were then matched to DGS signs with the same iconic motivation, and similar form. We explicitly constructed pairs that did not completely overlap in form, yet did also not completely lack form overlap. Figure 4 shows the high-overlap gesture-sign pair BICYCLE, in which only the size of the movement differs. In the low-overlap pair COMMUNICATION, the aperture of the fingers, position of the thumbs, location, movement type and alternation differ, but several features are the same, such as the selected fingers, handedness and repetition of movement (Figure 5).

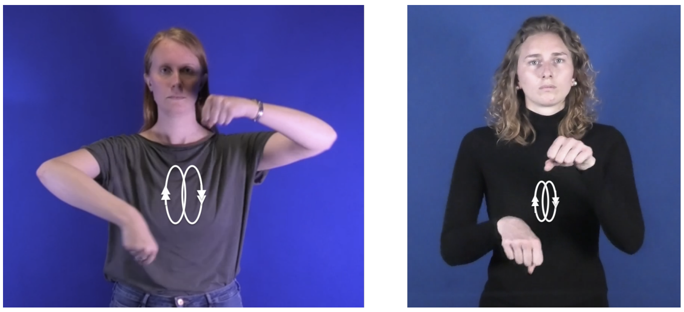

Figure 4: Gesture (left) and DGS sign (right) for BICYCLE.

3.2 Transparency and iconicity ratings

We used PsychoPy's (Peirce et al. 2019) online platform Pavlovia[3] to perform the experiment and Prolific[4] as the recruitment tool. 32 people participated in one of two versions of the pilot experiment: 16 participants saw a version with gestures; 16 saw a version with DGS signs. All participants were unfamiliar with any sign language, lived in Germany and reported German

3. https://pavlovia.org/
4. https://www.prolific.co/

Figure 5: Gesture (left) and DGS-sign (right) for COMMUNICATION.

as their L1. We elicited ratings from hearing sign-naive participants for two reasons. First, this population is the target of the next study these items will be used in, and as such, their judgements provide better insight into the perception of iconicity by the target population than those of a different population (i.e. (deaf) signers). Secondly, (deaf) signers may be influenced in their iconicity judgments by the lexical status of the items (gestures and signs), obfuscating the ratings of purely iconicity.

The task consisted of two parts, for which the full instructions can be found in the supplementary materials: https://osf.io/x4srz/. In both parts of the experiment the items (a subset of the eventual experiment items, totalling 32 pairs) were presented in a randomized order as continuous, looped videos, to eliminate the possibility that participants relied on their memory of the item rather than the item itself. In the first part, participants were asked to guess the meaning of the gesture/sign they were shown, and type out their German translation for it. They were instructed that there was no wrong answer, but were asked to use as few words as possible. In the second part, they were shown the same items, but were now also provided with the German translation, and asked to rate the iconicity of the item, as well as the iconicity of the individual parameters: the handshape, location and movement of the item. We instructed them to rate iconicity as ’how much a sign/gesture looks like what it means’ on a 7-point scale. The scale was, however, not numbered, and only the two extremes were labeled as ’not iconic’ and ’very iconic’ respectively.

The transparency score of a given item was determined by the proportion of participants that correctly guessed the meaning. For this, the German translations provided by the participants in the first part of the task were scored for their correctness. We opted for a scalar system instead of a binary system, as the additional information about the degree of (in)correctness can provide insight into exactly where the troubles in identifying a concept stem from. We adapted the scoring criteria of the Boston Naming Task (Roomer et al. 2011) and Ortega, Schiefner, and Özyürek (2019) to the following scoring system (Table 1). Correct answers, or answers that included the target concept received a score of 3. Oftentimes, answers were close, but not fully correct, e.g. participants omitted the target concept, or provided multiple answers (indicating uncertainty). These close, but not quite correct answers received a score of 2. A score of 1 was given to responses where there was some relation to the target, but the focus remained (too much) on visual features of the item, rather than abstracting it to the intended referent. The final category of 0 contains all answers that do not have any relation to the intended concept. To calculate the transparency score for an item, we took the number of items that received a score of 3, and divided that by the total number of responses. Items scored 0-2 were considered incorrect. The resulting score lies between 0 and 1, with 0 being no correct guesses and 1 being only correct guesses. Note that since we are scoring responses to both gestures and signs, we do not consider possible translations of an item other than the intended concept correct, as opposed to Trettenbrein and colleagues (2021). While 'drive' may be a correct translation for the DGS sign CAR , we cannot know whether the same would be applicable to the gesture for CAR . To keep the transparency scores comparable between the item types, these possible translations received a score of 2.

Table 1: Meaning guesses scoring scale.

Score	Criteria	Example: CAR
3 - Correct	(Includes) target concept (word class may vary).	'car', 'drive a car'
2 - Underspecified	• Does not (fully) include target concept, or; • Contains multiple answers.	'drive', 'steer' 'drive/car'
1 - Related	Relation to the target, but focus on a visual feature (without abstraction).	'turn something'
0 - Unrelated	No relation to the target concept.	'work'

3.3 Statistical analyses

Using R Statistical Software (v4.3.1; R Core Team 2023), we calculated the mean iconicity rating of the whole item, handshape, location and movement of each individual item, for both item types (gesture and sign). The transparency score is already a value averaged per item. We then calculated the means and standard deviations of these five measurements for the two item types. We expected the distribution of some of these not to conform to normality. Firstly, because we were dealing with pilot data and the sample is not as large (both in terms of participants and items). Secondly, since transparency scores are known to skew heavily towards the lower end of the scale (e.g. Sevcikova Sehyr and Emmorey 2019; Trettenbrein et al. 2021; Van Nispen, Van De Sandt-Koenderman, and Krahmer 2017). Therefore we employed non-parametric tests to explore relationships between variables. We used the Wilcoxon rank sum test to determine if signs and gestures meaningfully differ in iconicity as well as transparency, and the Kendall rank correlation test to quantify the correlation between whole-item iconicity and individual parameter iconicity.

4 Results and discussion

4.1 Overview

The mean and standard deviation of the dependent variables for both types of items (gesture and sign) can be found in Table 2. The transparency score is set on a 0 to 1 scale, with 0 being no correct meaning guesses, and 1 being only correct answers. The iconicity ratings are averaged item ratings, set on a scale from 1 (not iconic) to 7 (very iconic).

Figure 6 shows the density plots of the measurements. The distribution of handshape (c) and movement (e) iconicity were potentially normally distributed when taking item type (gesture/sign) into account, as was that of iconicity ratings of whole gestures (but not signs, b). The transparency scores (a) are heavily skewed towards the low end of the scale, whereas

Table 2: Mean and standard deviation of transparency score and iconicity ratings (whole-item, handshape, location and movement) of both gestures and signs.

	Gestures		Signs	
	Mean	SD	Mean	SD
Transparency score	0.30	0.35	0.19	0.26
Iconicity whole item	5.38	1.15	4.89	1.49
Iconicity HS	5.16	1.13	4.73	1.47
Iconicity LOC	5.32	1.34	4.87	1.66
Iconicity MOV	5.49	1.05	5.00	1.33

the iconicity of whole signs (b) and that of location (d) were skewed towards the higher end of the scale. Several of the measurements also displayed a bimodal distribution.

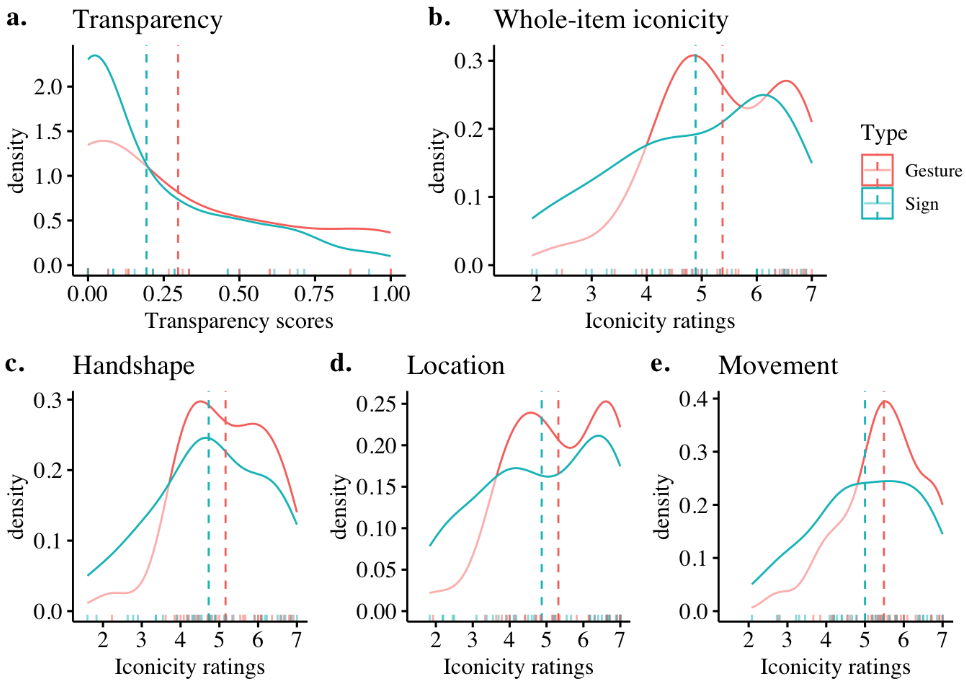

Figure 6: The density plots for transparency (a), whole-item iconicity (b), handshape iconicity (c), location iconicity (d) and movement iconicity (e) of gestures and signs.

4.2 Item type

The means of the transparency score and iconicity ratings are higher for gestures than signs (Figure 7). The median transparency score for gestures was 0.133 (IQR = 0.525), that for signs was 0.075 (IQR = 0.298). The Wilcoxon test showed this was not a significant difference (p = 0.316). Similarly, the difference in median iconicity rating for gestures (5.31, IQR = 1.8) and signs (5.04, IQR = 2.18) was not significantly different (p = 0.177).

On the basis of these results, we can not draw any conclusions about a difference in the transparency and perceived iconicity of gestures and signs. However, the occurrence

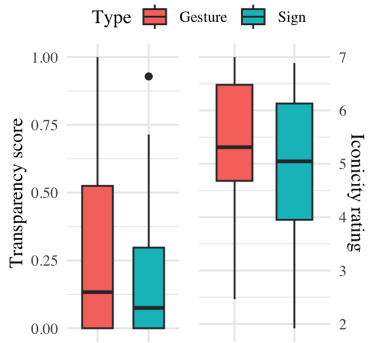

Figure 7: Transparency and iconicity by Type.

of higher iconicity ratings for signs in several pairs is noteworthy (Figure 8). The strongest example of this is MOTORBIKE (Figure 9), with a whole-item iconicity rating of 6.22 for the sign and 3.92 for the gesture. The form difference between the two is found in the handedness (two-handed symmetrical gesture, two-handed asymmetrical sign), orientation change (flexion and extension in the gesture, only extension in the sign) and repetition (repeated movement in the gesture, single movement in the sign). This and other 'counter' examples are an indication that gestures are not necessarily more iconic than signs.

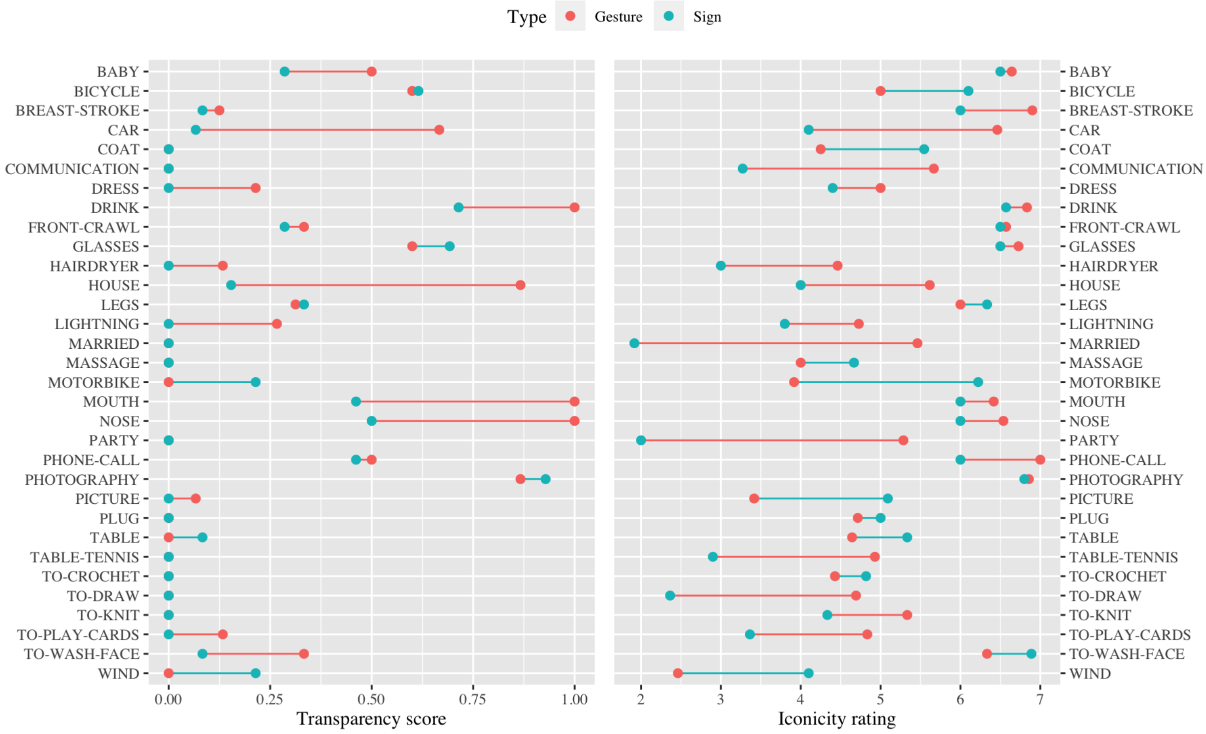

Figure 8: The average transparency scores and iconicity ratings of the concepts, separated by item type. Line coloring indicates the highest scored/rated type, single dots signify identical scores for both gesture and sign.

As such, the tendency for signs to lose iconicity over time through phonological changes, as described by Frishberg (1975) for ASL, may be more nuanced (or more nuanced for DGS). If we take the present-day gestures as a proxy for the gestural input at the introduction of the sign into DGS, we would expect the gesture to be rated as more iconic, and per Frishberg's

Figure 9: Gesture (left) and DGS sign (right) for MOTORBIKE.

first 'goal' of symmetry, we would also expect the sign to be more symmetrical than the gesture. Neither of these predictions hold, and in fact, the opposite is suggested. In our set of 32 gesture-sign pairs, 11 signs were found to be more iconic than their gesture counterparts (see Figure 8), and four signs (COMMUNICATION, MOTORBIKE, PHOTOGRAPHY, TO-CROCHET), contained some asymmetry that was not found in the gesture. The reverse (i.e. change towards symmetry) was only found in two signs (PARTY, TO-KNIT).

In all four signs that moved towards asymmetry, this may be explained by the iconicity of this asymmetry. In COMMUNICATION the alternating movement maps onto the alternation of conversational turns, and in MOTORBIKE, PHOTOGRAPHY and TO-CROCHET the movement of one hand instead of two maps onto the real world referent, where only one hand accelerates, clicks the shutter or moves the crochet hook. With a larger sample size and a more systematic investigation of this phenomenon, we may find evidence that gives us further insight into how changes in sign form over time relate to perceived iconicity.

4.3 Parameters

According to the Kendall rank correlation test, each of the individual parameters (handshape, location and movement), is significantly correlated with whole-item iconicity in both gestures and signs, with correlations ranging from 0.69 to 0.87 (Figure 10).

Between the individual items, there is variation in how the parameters are rated for iconicity, and moreover how they are ranked within an item (Figure 11). Comparing for example the concepts HAIRDRYER and LIGHTNING, both have one parameter rated much higher than the other two (in both gesture and sign), but in HAIRDRYER, this is the location, whereas in LIGHTNING, it is the movement. The different highest-rated parameter may be explained by the different iconic strategies these items employ. The acting strategy of handling a hairdryer allows for the iconic mapping of the location, and in the tracing strategy of LIGHTNING (tracing a prototypical shape of a lightning bolt in the air), movement is the parameter that conveys the most iconically mapped meaning. However, this may also be caused by the higher saliency of the parameters location and movement (Baus, Gutiérrez, and Carreiras 2014; Hildebrandt and Corina 2002). Handshape is the most highly rated parameter in only 9 (4 gestures, 5 signs) of the items, and never by a large margin. The combined use of parameter iconicity ratings and saliency measures may disentangle their effects.

4.4 Planned methodological changes

To decrease the influence of our methodology on the non-normal distributions, we will make a few changes in the actual experiment. Firstly, we will adopt an 11-point scale in the rating

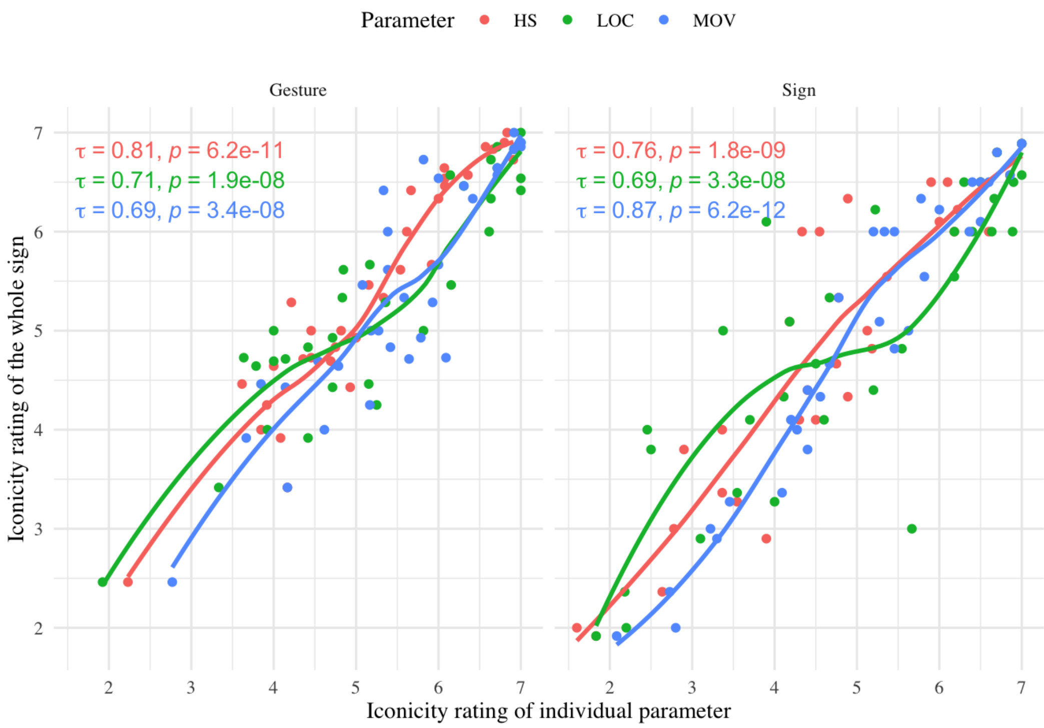

Figure 10: The correlations between the iconicity of the individual parameters and the whole-item.

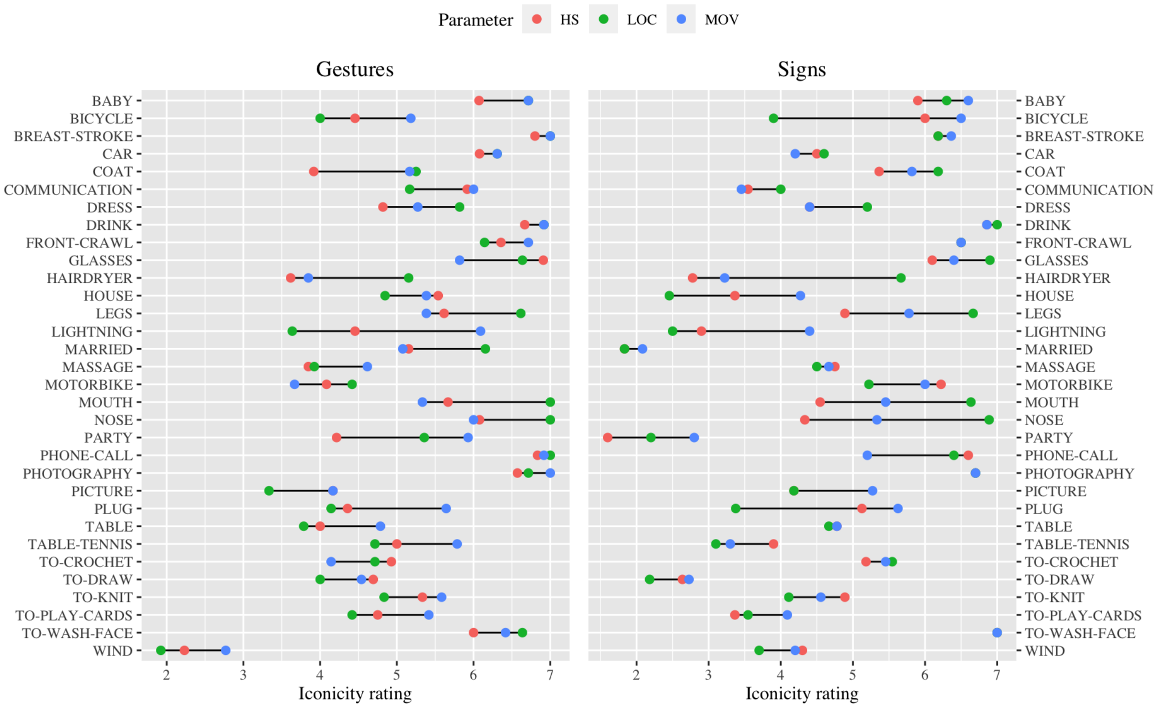

Figure 11: The average iconicity ratings of the parameters handshape, location and movement of the individual concepts.

of iconicity, to be able to reliably treat iconicity ratings as continuous variables in the statistical analysis (Wu and Leung 2017). Secondly, the non-normal distributions may be related to the relatively low number of data points, as well as to an unbalanced set of items. The experiment will include 100 participants and 120 items, which will drastically increase the number of data points. The items will also be more balanced for form overlap between gesture and sign (60 high-overlap, 60 low-overlap). In the items in the pilot the high-overlap category was over-represented (25/32 items), making an analysis of this variable impossible.

To improve our grasp of the quality of the ratings provided by participants, we will also include repeated trials in the experiment. This will allow us to catch strong discrepancies within a rater over the course of the experiment.

Additionally, as the majority of items are selected on the basis of having a gesture counterpart (i.e. a predominant gesture produced by participants), which skews the set towards iconicity and 'expressibility', it will likely not be a reflection of the full lexicon of DGS. To still provide participants with items that are possibly not very iconic, 20 fillers without a gestural counterpart will be added to the experiment. These fillers are signs for concepts that were elicited in the silent gesture task, received a 'pass' by at most one participant, but did not yield a predominant gesture.

5 Conclusion

On the continuum of conventionalization, there are multiple forces pushing and pulling gestures or signs toward higher or lower iconicity. On the one hand, the system a sign has to conform to may restrict the exact mapping of the referent onto the hands, but on the other hand, repeated use in interaction (and the feedback and meaning negotiation this entails, Byun et al. 2022) may make users of an (emerging) language converge on a form that is more 'true' to the referent. Furthermore, the individual parameters' iconicity is correlated to the whole-item iconicity, but is possibly also linked to parameter saliency and iconic strategy.

Our methodology of collecting iconicity ratings for both whole items and parameters successfully quantified the relative iconicity of (a) gestures and signs and (b) the parameters handshape, location and movement. Results from the pilot experiment reported here do not provide support for a difference in iconicity between gestures and signs, but do hint at a role for characteristics of individual gesture-sign pairs. They also show a correlation between parameter iconicity and whole-item iconicity. We aim to continue to improve this methodology and use it to gain a more fine-grained understanding of iconicity in the manual modality.

Acknowledgments

We acknowledge that the silent gesture data were generously provided by the project From Icon to Abstraction (Pamela Perniss, Gerardo Ortega and Annika Schiefner). We thank the reviewer for their helpful comments and suggestions. This paper is part of the project "The Gesture-to-Sign Trajectory: Phonological Parameters in Production and Real-Time Comprehension" (project number 501989071) of the Priority Program Visual Communication (ViCom), funded by the German Research Foundation (DFG SPP 2392 Visual Communication (ViCom)).

References

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

Baus, Cristina, Eva Gutiérrez, and Manuel Carreiras. 2014. "The Role of Syllables in Sign Language Production." *Frontiers in Psychology* 5 (November 13, 2014). https://doi.org/10.3389/fpsyg.2014.01254.

Byun, Kang-Suk, Seán G. Roberts, Connie De Vos, Ulrike Zeshan, and Stephen C. Levinson. 2022. "Distinguishing Selection Pressures in an Evolving Communication System: Evidence from Color-Naming in "Cross Signing"." *Frontiers in Communication* 7 (December 16, 2022): 1024340. https://doi.org/10.3389/fcomm.2022.1024340.

Cates, Deborah, Eva Gutiérrez, Sarah Hafer, Ryan Barrett, and David Corina. 2013. "Location, Location, Location." *Sign Language Studies* 13 (4): 433–461.

Dingemanse, Mark, Marcus Perlman, and Pamela Perniss. 2020. "Construals of Iconicity: Experimental Approaches to Form–Meaning Resemblances in Language." *Language and Cognition* 12, no. 1 (March): 1–14. https://doi.org/10.1017/langcog.2019.48.

Frishberg, Nancy. 1975. "Arbitrariness and Iconicity: Historical Change in American Sign Language." *Language* 51, no. 3 (September): 696. https://doi.org/10.2307/412894. JSTOR: 412894.

Goldin-Meadow, Susan, and Diane Brentari. 2017. "Gesture, Sign, and Language: The Coming of Age of Sign Language and Gesture Studies." *Behavioral and Brain Sciences* 40:e46. https://doi.org/10.1017/S0140525X15001247.

Hildebrandt, Ursula, and David Corina. 2002. "Phonological Similarity in American Sign Language." *Language and Cognitive Processes* 17, no. 6 (December): 593–612. https://doi.org/10.1080/01690960143000371.

McNeill, David. 2000. *Language and Gesture.* Vol. 2. Language Culture and Cognition. Cambridge: Cambridge University Press. ISBN: 978-0-521-77761-2.

Occhino, Corrine, Benjamin Anible, Erin Wilkinson, and Jill P. Morford. 2017. "Iconicity Is in the Eye of the Beholder: How Language Experience Affects Perceived Iconicity." *Gesture* 16, no. 1 (June 29, 2017): 100–126. https://doi.org/10.1075/gest.16.1.04occ.

Omardeen, Rehana. 2018. "The Contribution of Lexical Overlap to Perceived Iconicity in Foreign Signs." Master's thesis, Radboud University Nijmegen.

Ortega, Gerardo, and Gary Morgan. 2015. "Input Processing at First Exposure to a Sign Language." *Second Language Research* 31, no. 4 (October): 443–463. https://doi.org/10.1177/0267658315576822.

Ortega, Gerardo, and Aslı Özyürek. 2020. "Systematic Mappings between Semantic Categories and Types of Iconic Representations in the Manual Modality: A Normed Database of Silent Gesture." *Behavior Research Methods* 52, no. 1 (February): 51–67. https://doi.org/10.3758/s13428-019-01204-6.

Ortega, Gerardo, Aslı Özyürek, and David Peeters. 2020. "Iconic Gestures Serve as Manual Cognates in Hearing Second Language Learners of a Sign Language: An ERP Study." *Journal of Experimental Psychology: Learning, Memory, and Cognition* 46, no. 3 (March): 403–415. https://doi.org/10.1037/xlm0000729.

Ortega, Gerardo, Annika Schiefner, and Aslı Özyürek. 2019. "Hearing Non-Signers Use Their Gestures to Predict Iconic Form-Meaning Mappings at First Exposure to Signs." *Cognition* 191 (October): 103996. https://doi.org/10.1016/j.cognition.2019.06.008.

Peirce, Jonathan, Jeremy R. Gray, Sol Simpson, Michael MacAskill, Richard Höchenberger, Hiroyuki Sogo, Erik Kastman, and Jonas Kristoffer Lindeløv. 2019. "PsychoPy2: Experiments in Behavior Made Easy." *Behavior Research Methods* 51, no. 1 (February 1, 2019): 195–203. https://doi.org/10.3758/s13428-018-01193-y.

Perniss, Pamela, Gabriella Vigliocco, and David P. Vinson. 2014. "The Bridge of Iconicity: From a World of Experience to the Experience of Language." *Philosophical Transactions of the Royal Society B: Biological Sciences* 369, no. 1651 (September 19, 2014). https://doi.org/http://dx.doi.org/10.1098/rstb.2013.0300.

Pietrandrea, Paola. 2002. "Iconicity and Arbitrariness in Italian Sign Language." *Sign Language Studies* 2 (3): 296–321. https://doi.org/10.1353/sls.2002.0012.

Poggi, Isabella. 2008. "Iconicity in Different Types of Gestures." *Gesture* 8, no. 1 (May 15, 2008): 45–61. https://doi.org/10.1075/gest.8.1.05pog.

R Core Team. 2023. *R: A Language and Environment for Statistical Computing.* Vienna, Austria: R Foundation for Statistical Computing. https://www.R-project.org/.

Roomer, E K, S Brok, A C Hoogerwerf, and D E Linn. 2011. *Een test voor woordvinding.*

Sandler, Wendy. 2011. *Phonological Representation of the Sign: Linearity and Nonlinearity in American Sign Language.* Vol. 32. Walter de Gruyter.

Schiefner, Annika, Pamela Perniss, and Gerardo Ortega. 2023. "Iconicity in Concrete and Abstract Concepts: A Cross-Linguistic Comparison of Silent Gestures and Signs." GeSpIn. Nijmegen. https://osf.io/8s4je.

———. Forthcoming.

Sevcikova Sehyr, Zed, and Karen Emmorey. 2019. "The Perceived Mapping between Form and Meaning in American Sign Language Depends on Linguistic Knowledge and Task: Evidence from Iconicity and Transparency Judgments." *Language and Cognition* 11, no. 2 (June): 208–234. https://doi.org/10.1017/langcog.2019.18.

Tkachman, Oksana. 2023. "Embodiment and Emergent Phonology in the Visual-Manual Modality: Factors Enabling Sublexical Componentiality." PhD diss., University of British Columbia.

Trettenbrein, Patrick C., Nina-Kristin Pendzich, Jens-Michael Cramer, Markus Steinbach, and Emiliano Zaccarella. 2021. "Psycholinguistic Norms for More than 300 Lexical Signs in German Sign Language (DGS)." *Behavior Research Methods* 53, no. 5 (October): 1817–1832. https://doi.org/10.3758/s13428-020-01524-y.

Van Nispen, Karin, W. Mieke. E. Van De Sandt-Koenderman, and Emiel Krahmer. 2017. "Production and Comprehension of Pantomimes Used to Depict Objects." *Frontiers in Psychology* 8 (July 11, 2017): 1095. https://doi.org/10.3389/fpsyg.2017.01095.

Wu, Huiping, and Shing-On Leung. 2017. “Can Likert Scales Be Treated as Interval Scales?—A Simulation Study.” *Journal of Social Service Research* 43, no. 4 (August 8, 2017): 527–532. https://doi.org/10.1080/01488376.2017.1329775.