Iconic Locations in Swedish Sign Language: Mapping Form to Meaning with Lexical Databases

Carl Börstell & Robert Östling
Department of Linguistics
Stockholm University
{calle,robert}@ling.su.se

Abstract

In this paper, we describe a method for mapping the phonological feature `location` of Swedish Sign Language (SSL) signs to the meanings in the Swedish semantic dictionary SALDO. By doing so, we observe clear differences in the distribution of meanings associated with different locations on the body. The prominence of certain locations for specific meanings clearly point to iconic mappings between form and meaning in the lexicon of SSL, which pinpoints modality-specific properties of the visual modality.

1 Introduction

1.1 Language and iconicity

The word forms of a language have traditionally been regarded as arbitrary, that is, there is no motivation for why a certain meaning is encoded by a specific form (de Saussure, 1916). The iconicity found in the word forms of spoken language is normally restricted to a few categories—e.g. onomatopoeia and ideophones (Perniss et al., 2010)—but also visible in so-called *phonaesthemes*, grouping certain meanings together—e.g. *tw-* in *twist* and *twirl* (Kwon and Round, 2015). Large-scale cross-linguistc comparisons of form and meaning have shown that there are some preferences for using and avoiding certain sounds for certain meanings (Blasi et al., 2016). However, since the extent of iconicity in spoken language is still quite limited, the general assumption is still that arbitrary word forms are the norm for any given language in that modality.

1.2 Signed language and iconic locations

Signed language uses the other of the two natural modalities of human language, being visual–gestural instead of auditive–oral. A key difference between signed and spoken language is that the former is widely regarded as more iconic (and consequently less arbitrary) than the latter, in terms of both lexically specified and morphologically modified depiction (Klima and Bellugi, 1979). The articulation of any sign is located in the physical space on or around the body of the signer. The location of the sign (a.k.a. *place of articulation*) can be iconic already in lexical signs (Taub, 2001), but sign locations may be altered to adhere to and syntax/discourse iconicity (Perniss, 2012; Meir et al., 2013).[1] In this study, we only focus on lexically specified locations of signs (see Section 2.1). Two examples of iconic locations in SSL signs are illustrated in Figure 1, in which the sign THINK is located at the forehead (representing brain activity), and Figure 2, in which the sign QUIET is located at the mouth (represented by a well-known gesture, depicting an obstacle in front of the lips).

Figure 1: The SSL sign THINK (Svenskt teckenspråkslexikon, 2016).

The iconic relationship between form and meaning is well-attested for signed language, including `location` as one form feature. However, few studies that have investigated this link by quantitative means, and none for SSL.

[1] The co-speech gestures often accompanying spoken language may be similarly iconic, for instance with regard to the location of gesturing in the physical space (McNeill, 1992).

Proceedings of the 21st Nordic Conference of Computational Linguistics, pages 221–225,
Gothenburg, Sweden, 23-24 May 2017. ©2017 Linköping University Electronic Press

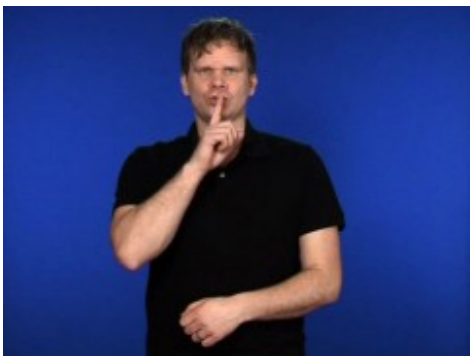


Figure 2: The SSL sign QUIET (Svenskt teckenspråkslexikon, 2016).

2 Data and Methodology

2.1 The SSL online dictionary

The SSL dictionary (SSLD) (Svenskt teckenspråkslexikon, 2016) is an online video dictionary of SSL. It is an ongoing language resource and documentation project, creating a lexical database constantly expanding in size (Mesch et al., 2012). The version used for this study included 15,874 sign entries. Each sign entry has one or more Swedish word translations, and also features a phonological transcription of the sign form, in which sign location is one value.

All sign data were exported from the SSLD database, and from this raw data, Swedish keywords and sign locations were extracted using a Python script. For the purposes of this study, complex signs with more than one location (e.g. compounds) were excluded.

For single location signs, we also excluded a) signs using the so-called *neutral space* as the location, and b) signs for which the other, non-dominant, hand was used as the location (Crasborn, 2011). The former were excluded since we were only interested in signs with body-specified locations.[2] The latter cases were excluded since the other hand is found to be iconic in terms of its shape and interaction with the dominant hand, rather than as a location *per se* (Lepic et al., 2016).

The finalized SSLD data consist of a list of 3,675 signs that met our criteria, their Swedish keywords, and location. In this list, 29 locations were present. These were collapsed into 20 locations, conflating near identical locations (e.g. `eyes` and `eye`). Table 1 shows a list of all locations and the number of signs per location.

[2]This does not necessarily entail body contact.

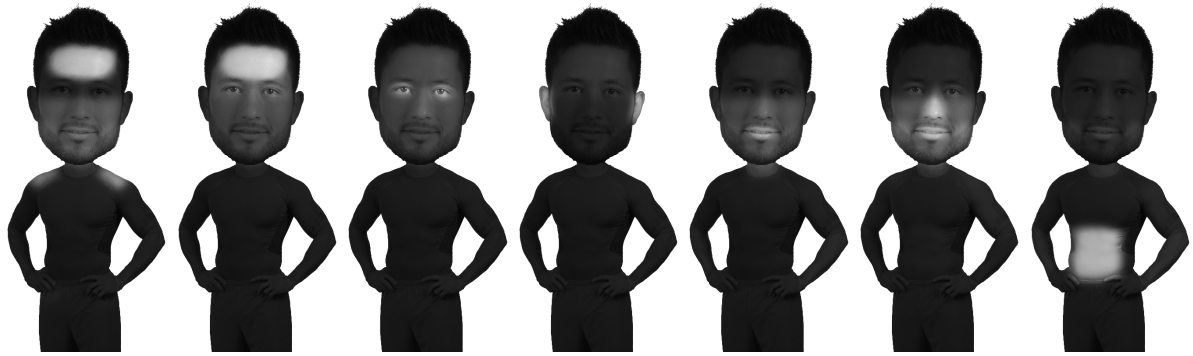
Location	No. of signs
head	81
forehead	414
upper face	159
eyes	95
face	153
nose	214
ears	103
lower face	47
cheeks	210
mouth	398
chin	325
neck	196
shoulders	77
arm	36
upper arm	47
lower arm	110
chest	860
belly	101
hip	42
leg	7
Total	3,675

Table 1: Distribution of signs across locations (anatomically descending).

2.2 SALDO

SALDO (Borin and Forsberg, 2009) is a semantic lexicon of Swedish, in which each word sense is arranged into a hierarchy through its (unique) *primary descriptor* and its (one or more) *secondary descriptors*. Unlike the more familiar WordNet (Miller, 1995) style lexica, the precise semantic relationship indicated by SALDO's descriptors is not formally specified. While this makes some of the applications of WordNet difficult to reproduce with SALDO, generating a number of broad semantic categories is sufficient for our needs.

For the purposes of this work, we define the semantic category defined by a word sense to be the set of all primary or secondary descendants in SALDO. This implies that each sense in SALDO defines a category, possibly overlapping, and that the choice of which categories to investigate is very free. We selected categories that were large enough to provide a sensible analysis, as well as semantically tied to the human body. Because SSLD does not contain any mapping to SALDO's word senses, we approximate sense disambiguation by using the first SALDO sense of any SSLD entry. In practice, this amounts to looking up the



(a) 'believe' (b) 'think' (c) 'see' (d) 'hear' (e) 'say' (f) 'feel' (g) 'eat'

Figure 3: Location distributions for seven semantic categories. Brightness represents the degree to which a given body part is over-represented in the given semantic category, with respect to the distribution over locations for *all* signs in the lexicon.

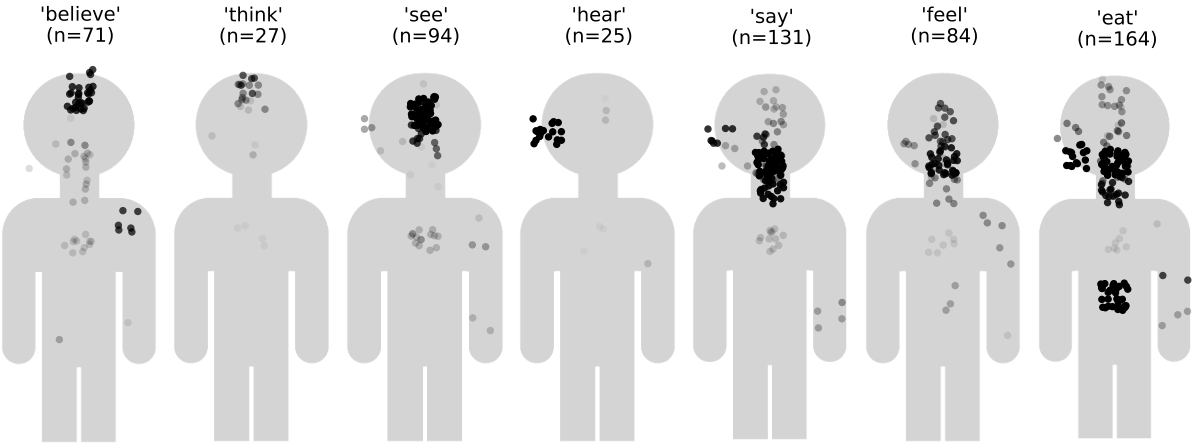


Figure 4: The distribution of locations for signs within seven semantic categories (with number of sign entries per semantic category in brackets).

Swedish translation available in each SSLD entry using SALDO, and choosing the first sense in case there are several. This is a surprisingly close approximation, because the first sense is generally the most common.[3]

To give a sense of how one of the semantic categories we study looks, we sample ten random signs in the category 'eat': *animal feed*, *appendix* (anatomy), *kiwi*, *gravy*, *foodstuff*, *lunch*, *belly ache*, *anorexia*, *full*, *oatmeal*. While many actual types of food are included, we also see terms such as *appendix* whose assocation to 'eat' is more indirect.

[3]The exception to this among our concepts is 'feel', where we use the second SALDO sense of the corresponding Swedish word, 'känna'.

2.3 Visualization

We investigate the distribution of locations for a given semantic category by first looking up its members in SALDO as described above, then looking up the corresponding signs in SSLD through their Swedish translations. The locations of the resulting set of signs is then visualized in two ways:

- by varying the light level of body parts proportional to the (exponentiated) pointwise mutual information (PMI) of the given concept and that location (see Figure 3).
- by a jitter plot showing the number of signs within a concept with a certain location (see Figure 4).

Pointwise mutual information is defined as

$$\mathrm{PMI}(l,c) = \log \frac{p(l,c)}{p(l)p(c)}$$

where, as we use maximum-likelihood estimation, $p(l)$ is the proportion of signs articulated at location l, $p(c)$ is the proportion of signs that belong to category c, and $p(l,c)$ the proportion that are both of the above at the same time. Intuitively, this is a measure of how overrepresented a location is among the signs within a given concept, relative to the overall distribution of locations in the SSLD lexicon. In our visualization, high PMI is represented by brighter regions.

We have chosen to use two separate but similar visualization techniques for reasons of clarity, since the first gives an intuitive picture of where on the body a particular semantic category is focused in SSL vocabulary, whereas the second makes it easier to see the actual distribution of sign locations within a concept without comparison to the overall distribution.

3 Results

Figure 3 shows the location distributions for seven semantic categories: 'believe', 'think', 'see', 'hear', 'say', 'feel', and 'eat'.

The amount of iconicity in SSL is clearly visible in this figure, where signs in the categories 'believe' and 'think' are over-represented around the forehead (with specific meanings such as *suspect* and *ponder*), 'see' around the eyes (e.g. *stare*), 'hear' on the ears (e.g. *listen*), 'say' around the mouth (e.g. *speak*, *talk*) or neck (e.g. *voice*), 'feel' on several locations on the lower face related to sensory inputs (e.g. *smell*, *sweet*), and 'eat' around the mouth (e.g. *lunch*) or belly (e.g. *hungry*).

This iconicity is by no means absolute, as indicated by Figure 4. This shows that even in the most extreme cases, such as 'hear' and 'think', the bias in location is not absolute. Other categories, like 'say', are in fact distributed quite widely throughout the body although the mouth area is clearly over-represented.[4]

4 Conclusions

In this paper, we have showed clear examples of iconic patterning in the distribution of meanings across the lexically specified locations of SSL signs. This is done by quantitative means, using a novel method of matching Swedish word entries in the SSLD to the meanings in the semantic dictionary SALDO, followed by a visualization based on a prominence-ranking of locations to meaning domains. The results illustrate that some body locations are much more prominent than others within certain semantic domains. This is attributed to the iconic structure of signed language, with sign forms directly or metaphorically evoking salient properties of some referent. Since not all signs are necessarily iconic, and because iconic forms may choose from a range of features of its referent to depict, the distribution of meanings to locations is not absolute. Instead, locations are more or less prominent for certain meanings, and in many cases this is directly linked to iconicity.

[4] In Figure 4, the prominence of each location is shown by level of darkness in the plotted signs (i.e. darker means more prominent).

Acknowledgments

We wish to thank the two anonymous reviewers for comments and suggestions on this paper.

References

Damián E. Blasi, Søren Wichmann, Harald Hammarström, Peter F. Stadler, and Morten H. Christiansen. 2016. Sound–meaning association biases evidenced across thousands of languages. *Proceedings of the National Academy of Sciences*, 113(39):10818–10823.

Lars Borin and Markus Forsberg. 2009. All in the family: A comparison of SALDO and WordNet. In *NODALIDA 2009 Workshop on WordNets and other Lexical Semantic Resources – between Lexical Semantics, Lexicography, Terminology and Formal Ontologies*, pages 7–12, Odense, Denmark.

Onno Crasborn. 2011. The other hand in sign language phonology. In Marc van Oostendorp, Colin J. Ewen, Elizabeth Hume, and Keren Rice, editors, *The Blackwell companion to phonology, vol. 1*, chapter 10, pages 223–240. Malden, MA & Oxford.

Ferdinand de Saussure. 1916. *Cours de linguistique générale*. Payot, Paris.

Edward S. Klima and Ursula Bellugi. 1979. Iconicity in signs and signing. In Edward S. Klima and Ursula Bellugi, editors, *The signs of language*, pages 9–34. Harvard University Press, Cambridge, MA.

Nahyun Kwon and Erich R. Round. 2015. Phonaesthemes in morphological theory. *Morphology*, 25(1):1–27.

Ryan Lepic, Carl Börstell, Gal Belsitzman, and Wendy Sandler. 2016. Taking meaning in hand: Iconic motivations for two-handed signs. *Sign Language & Linguistics*, 19(1):37–81.

David McNeill. 1992. *Hand and Mind: What Gestures Reveal about Thought*. University of Chicago Press, Chicago, IL.

Irit Meir, Carol Padden, Mark Aronoff, and Wendy Sandler. 2013. Competing iconicities in the structure of languages. *Cognitive Linguistics*, 24(2):309–343.

Johanna Mesch, Lars Wallin, and Thomas Björkstrand. 2012. Sign Language Resources in Sweden: Dictionary and Corpus. In Onno Crasborn, Eleni Efthimiou, Evita Fotinea, Thomas Hanke, Jette Kristoffersen, and Johanna Mesch, editors, *Proceedings of the 5th Workshop on the Representation and Processing of Sign Languages: Interactions between Corpus and Lexicon [LREC]*, pages 127–130, Paris. ELRA.

George A. Miller. 1995. WordNet: A lexical database for English. *Communications of the ACM*, 38(11):39–41.

Pamela Perniss, Robin L. Thompson, and Gabriella Vigliocco. 2010. Iconicity as a general property of language: evidence from spoken and signed languages. *Frontiers in Psychology*, 1(227).

Pamela Perniss. 2012. Use of sign space. In Roland Pfau, Markus Steinbach, and Bencie Woll, editors, *Sign language: An international handbook*, pages 412–431. De Gruyter Mouton, Berlin/Boston, MA.

Svenskt teckenspråkslexikon. 2016. Sign Language Section, Department of Linguistics, Stockholm University. http://teckenspraklexikon.ling.su.se/.

Sarah F. Taub. 2001. *Language from the body: Iconicity and metaphor in ASL*. Cambridge University Press, Cambridge.