Semi-automatic Extraction of Handshape Inventory in Russian Sign Language

Anna Klezovich

Sign languages are a very rewarding field for researching phonological structure and building blocks of language due to the visual modality. Although there are a number of formal phonological theories for sign languages, such as the Movement-Hold model (Liddell & Johnson 1994), the Prosodic model of phonology (Brentari 1998), etc., there is not much information on phonological inventories in various sign languages, so it would be fruitful to add data from more sign languages into the discussion. Furthermore, all previous research on the topic (as far as the author knows) used solely elicitation and manual annotation as methods, e.g. (van der Kooij 2002), (Nyst 2007), (Marsaja 2008), and (Schuit 2014). So, the aim of this research is: (1) to develop a semi-automatic method of handshapes extraction; (2) and to describe phonetic and phonemic inventories of one of the phonological features (Stokoe 1960; Battison 1978), namely, handshape, in Russian Sign Language (RSL).

The algorithm developed in this work is based on the Börstell's (2018) script for extracting hold position from videos to make sign overlay pictures of NGT and SSL signs, but modified to be applicable to RSL data. The idea is that in order to get an inventory of phonetic handshapes of a language it is sufficient to annotate them only in the positions without movement (aka. holds) (Liddell & Johnson 1994). In this work I use videos of all RSL signs from Spreadthesign online dictionary (Hilzensauer et al. 2015) with phrases and compounds removed (3727 videos). For each video my algorithm cuts out frames without a moving object (i.e. hand(s)), then it calculates a histogram of color for each remaining frame, and, finally, it successively calculates differences between pairs of adjacent frames. Consequently, I treat each video as a signal, or to be more specific as a relation between the difference between two consecutive frames' histograms of color and frame numbers (roughly speaking – time). Finally, each signal is smoothed with the help of moving average to make prominent peaks stand out (see Fig. 1). In the resulting signal, negative peaks correspond to positions without movement, and the algorithm takes one or at most two highest negative peaks from each signal and returns snapshots of the corresponding frames with the holds. The manually established accuracy of this algorithm on my dataset is 76.7%. The core of the algorithm itself is significantly different from the script in (Börstell 2018), since it uses moving average instead of continuous wavelet transform and is more interpretable.

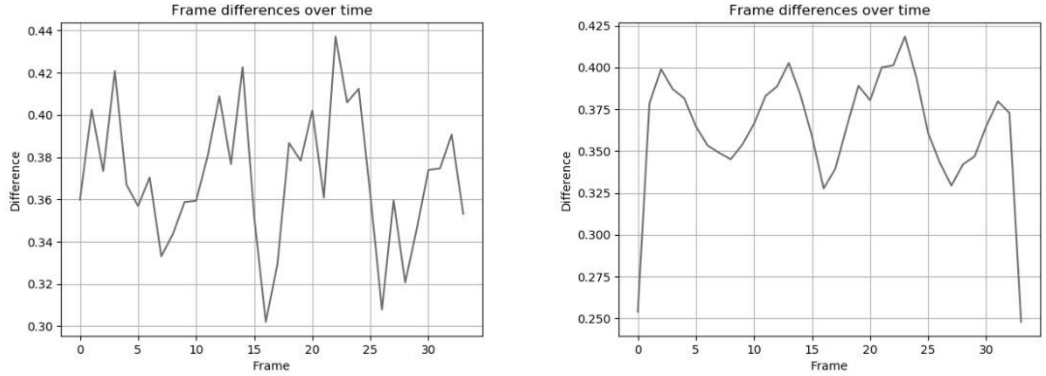
After automatic holds extraction, I manually annotated handshapes in all 5189 holds using the Hamburg Notation System (Hanke 2004). As a result, I obtained a list of 115 phonetic handshapes, which is rather typologically unusual. For instance, Schuit (2014) proposed that it is unlikely to find a sign language with more than 80 phonetic handshapes.

Finally, I preliminary established which handshapes could be allophones of each other, using van der Kooij's (2002) model. This phonological model proposes that, apart from taking into account individual differences between signers, handshapes should not be treated as phonological if they can be explained by iconicity or the ease of articulation. Using solely these two criteria, I inferred that RSL has 23 phonemic handshapes; however, the individual differences are not taken into account due to the nature of the dataset.

All in all, this work (1) proposes a new algorithm for holds extraction, which can potentially speed up the process of annotation for other sign languages when calibrated on new data, and (2) describes phonetic and phonemic inventories of handshapes on an extensive dataset in RSL. In the future, I am going to add orientation and location features to the analysis, so that I can explore (e.g. with the help of dimensionality reduction techniques) whether there are any hierarchical relationships between these three features, as the Prosodic model of phonology suggests (Brentari 1998).

Appendices:

Figure 1. FALL-IN-LOVE (RSL) sign as a signal. Left – before smoothing, right – after smoothing.



References:

Battison, Robbin (1978). Lexical borrowing in American sign language. Silver Spring, MD: Linstok Press.

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

Börstell, Carl (2018). Make signs still. (https://github.com/borstell/make_sign_stills)

Hanke, Thomas (2004). HamNoSys - representing sign language data in language resources and language processing contexts. In: O. Streiter & C. Vettori (eds.), Proceedings of the Workshop on Representation and Processing of Sign Languages. Paris: ELRA, 2004, 1–6.

Hilzensauer, Marlene and Klaudia Krammer (2015). A multilingual dictionary for sign languages: "spreadthesign". (https://www.spreadthesign.com/)

Kooij, Els van der (2002). Phonological categories in sign language of the Netherlands: the role of phonetic implementation and iconicity.

Liddell, Scott & Robert E. Johnson (1994). American Sign Language: The Phonological Base. Sign Language Studies, 64, 195–278.

Marsaja, I. Gede (2008). Desa Kolok. A deaf village and its sign language in Bali, Indonesia. Nijmegen: Ishara Press.

Nyst, Victoria (2007). A Descriptive Analysis of Adamorobe Sign Language (Ghana). PhD Dissertation, University of Amsterdam.

Schuit, Joke (2014). Signs of the arctic: Typological aspects of Inuit Sign Language. PhD Dissertation, University of Amsterdam.

Stokoe, William (1960/2005). Sign Language Structure: An Outline of the Visual Communication Systems of the American Deaf. Journal of Deaf Studies and Deaf Education, 10(1), 3–37.