Measuring phonological complexity in Sign Languages

[1]Valentina Aristodemo, [1]Chiara Annucci, [1,3]Justine Mertz, [2]Giustolisi Beastrice, [3]Carlo Geraci and [1]Donati Caterina
[1]University of Paris-Diderot, [2]University of Milano-Bicocca, [3]CNRS Institut Jean-Nicod

Goals. We propose a new data-driven measure of sign language (SL) perceptual/articulatory complexity. If satisfactory, the measure will be used to assess in parallel with or in absence of other measures (e.g., frequency and acquisition). We compare the performance of our data-driven measure with a theory-driven measure of complexity based on the SL feature geometry (Brentari, 1998). Our empirical base is composed of four sign languages: French, Italian, Catalan and Spanish (LSF, LIS, LSC and LSE).

Background. There are two main types of approaches to phonological complexity in spoken languages, which we can refer to *theory-driven* and *data-driven*. The *theory-driven approach* is well illustrated by Clements (1985) and Sagey (1986) where counting distinctive features is the relevant measure. *Data-driven approaches* typically rely on pattern/order of acquisition, frequency, speech errors and similar measurable facts to assess phonological complexity. Ideally the two measures should converge.

As for SL *theory-driven approaches* to complexity, Brentari's Prosodic Model (1998) belongs to the former tradition: each phonemic class is assigned a set of features organized in a hierarchical/geometric structure. Each sign can be described in terms of a branching tree. The root corresponds to the prosodic word and branches correspond to the phonemic classes of handshape, location, and movement, each containing its own feature geometry. The richer is the structure (in terms of positively specified features), the higher is the complexity of a sign. While this model provides important crosslinguistic generalizations, its validity beyond ASL cannot be taken for granted. Similar considerations hold for other models (e.g., Sandler & Lillo-Martin 2006 and van Der Kooij 2002).

As for *data-driven approaches*, frequency and acquisition data are available only for a few sign languages (e.g., ASL, BSL, NGT), but they are entirely missing for others and would require long-term efforts to obtain. Diagnostics based on error rates can provide a quick and handy measure for early detection of a number language disorders.

Definitions of complexity. *Our data-driven measure* is based on error rates in naïve non-signers. The rationale is: signs that can be accurately and fluently repeated are treated as simple (see below). *Theory-driven measure*. Adapting Brentari's model, we measured complexity by counting the number of nodes and features necessary to describe it.

Methods. **Data-driven measure**. A repetition task is used to assess sign complexity in non-signers. The procedure is identical for all SLs in the study. We describe here the case of LSF.

Materials. 108 signs in LSF were selected based on criteria such as lack of major iconicity, frequency, lack of regional variation. A Deaf consultant was video-recorded producing the citation forms of the signs.

Participants. 20 hearing non-signers acquainted with the visual culture of France were recruited (divided in 5 age groups: 18-29: 30-40; 40-49; 50-59; 60-70).

Task: Each participant was asked to watch once the video of a sign and (try to) repeat it. Their performance was video-recorded (2160 tokens).

Coding complexity. Two students with a basic competence in LSF coded the data according to 5 measures: fluency, accuracy in handshape, orientation, location and movement. For each measure we assigned 1 if correct, 0 if incorrect. Overall accuracy for a sign is obtained by summing accuracies in each component. The degree of accuracy is directly mapped onto a complexity scale (5 = all correct = least complex, 0 = all incorrect = most complex).

Theory-driven measure: A portion of the entire dataset is used as a pilot study. *Materials*. We annotated 15 items out of 108 used in the data-driven measure. These 15 items received a variable level of complexity in the data-driven measure (5 have a high level of complexity, 5 a low level of complexity and 5 have an intermediate level of complexity).
Coding complexity. The level of complexity depends on the number of nodes and positively specified features in its representation (lower values = less complex, higher values = more complex signs). The total set of nodes and features considered is 116 (handshape=67, location=22, movement =27).

Results. We report here preliminary results form LSF. By the time of the conference, the LSF and LIS data will be available for comparison.

Data-driven method. The overall mean of accuracy is 4.282 (SD=0.82). The most complex sign is HEDGEHOG with an average score of 3.15, while the easiest sign is HAM with a score of 5. Handshape is the class in which most of the errors are observed (45%) followed by movement (39%), orientation (8%) and location (7%). A mixed-model analysis with item and participant as random factors was conducted. A significant main effect of *age* was found (p=0.03). Younger participants are more accurate.

Theory-driven method: our 15 stimuli have an index of complexity that range from 15 to 39. The simplest sign is BAND-AID, while the most complex one is PEN.

Analysis. We observe a correlation between the overall complexity of the theory-driven measure and the overall accuracy/complexity of the data-driven method (r=-0.30). However, it is not significant (p=0.28). We observe a significant correlation between handshape complexity in theory-driven measure and overall accuracy (r=-0.58; p=0.02). The higher the complexity in the handshape is, the lower is the level of accuracy. Other correlations are not significant (movement/location vs. overall accuracy).

Discussion. Preliminary results show that: 1) The two measures converge; 2) Handshape is the class that better correlates with overall accuracy; 3) Still, for some signs we observed considerable divergence. The table illustrates that there are signs that receive a low score in complexity as measured by the theoretical model (i.e., that are predicted to be simple), but still have a poor performance in overall accuracy (e.g., SAUCE), and vice versa (e.g., BONE).

Data-driven scale		**Theory-driven scale**	
FLOWER	4,9	BAND_AID	16
RED	4,7	SAUCE	21
BAND_AID	4,55	TREE	25
PEAR	4,55	CASTLE	26
BONE	4,5	RED	26
TREE	4,4	PEAR	27
BREAD	4,35	GLASS	28
GLASS	4,2	FLOWER	29
CASTLE	4,1	COMPASS	32
LEAF	3,95	BREAD	32
PEN	3,9	THEATRE	34
FACTORY	3,9	LEAF	35
THEATRE	3,85	BONE	35
SAUCE	3,6	FACTORY	36
COMPASS	3,35	PEN	39

We shall speculate on the source of this divergence. In principle, this could be due to at least one of the following reasons: a) the data-driven measure, being non-linguistic, does not capture some important phonological categorizations; b) the theory-driven measured is not fully equipped to predict complexity in LSF. To address these issues, one could replicate this study in two ways: with signers as participants, and by using pseudo-signs as stimuli. We also expect comparison with the results of the study in LIS to shed light on these issues. Another interesting issue is whether handshape alone is enough to predict complexity. If this is the case, what is the role of place of articulation and movement in determining complexity? One possibility is that location and movement require a fully-fledged phonology in place. In this case, we expect major differences between signers and non-signers.

Selected References: **Brentari**, D. (1998). *A prosodic model of sign language phonology*. MIT Press. **Clements**, G. N. (1985). The geometry of phonological features. *Phonology Yearbook* 2, 225-252. **Sagey**, E. (1986). The representation of features and relations in nonlinear phonology. Doctoral dissertation, MIT, Cambridge, Mass.