# ON HOW TO MEASURE PHONOLOGICAL COMPLEXITY OF SIGN LANGUAGES

*Valentina Aristodemo[1], Chiara Annucci[1], Carlo Geraci[2], Beatrice Giustolisi[3], Doriane Gras[1], Justine Mertz[1,2] and Caterina Donati[1]*

[1] Université Paris Diderot, [2]Institu Jean Nicod (CNRS-ENS-EHESS), [3] Univeristà Milano Bicocca

**Introduction.** Having a measure of the phonological complexity of signs is an important challenge for anyone interested in developing assessment tests for evaluating language competences and language impairments in Deaf populations. Traditionally, there are two main types of approaches to phonological complexity, which we can refer to as theory-driven and data-driven, respectively. The theory-driven approach is well illustrated by Clements (1985) and Sagey (1986) where counting distinctive features is the relevant measure. Brentari's Prosodic Model (1998) belongs to this tradition: she assigns to each phonemic class a set of features organized in a hierarchical structure. Each sign can be described in terms of a branching tree where each branch corresponds to a phonemic class (handshape, location, and movement) and each node corresponds to a set of features. The richer the structure, the higher the complexity of a sign. While this model has proven to provide important cross-linguistic generalizations, its validity beyond ASL cannot be taken for granted.

Data-driven approaches, on the other hand, typically rely on pattern/order of acquisition, frequency, speech errors and similar measurable facts to assess phonological complexity. While some of these data are available for some sign languages (typically ASL and BSL), they are entirely missing for others.

**Goals.** We explore a new data-driven measure of complexity based on perceptual/articulatory criteria. If satisfactory, the measure can be employed for sign languages for which frequency and acquisition data are unavailable. We then compare this data-driven measure of complexity with a theory-driven measure 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).

**Definition of complexity.** *Data-driven measure*. We consider as simple those signs that can be accurately and fluently repeated by non-signers (see below for the coding).

*Theory-driven measure*. We adapted Brentari's model and measured sign 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 while uttering these signs in isolation.

*Participants*. 20 hearing non-signers acquainted with the visual culture of France were then 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 videorecorded (2160 tokens).

*Coding complexity.* Two students with a basic competence in LSF coded each repetition according to: fluency, accuracy in handshape, orientation, location and movement. To these components a binary value was assigned (correct=1; wrong=0). For each sign, the overall accuracy was obtained by summing the value of accuracy of these 5 components. The degree of accuracy is directly mapped onto a complexity scale (5 = least complex, 0 = most complex).

**Theory-driven measure**: A portion of the entire dataset was 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 is 116 (handshape=67, location=22, movement =27).

**Results**. As of today, we are able to report only on the first application of this protocol to LSF. By the time of the conference, the number of items analysed and compared wrt the two methods will have increased.
*Data-driven method.* The overall mean 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 subject as random effects and age as fixed effect 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.

**Assessing the convergence.** The correlation between theory- and data-driven measures on overall complexity and each class of phonemes was calculated. We observe a correlation between the overall complexity of the theory-driven measure and the overall accuracy 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 some interesting facts: 1) The two measures indeed converge; 2) Handshape is the class that better correlates with overall accuracy; 3) Still, for some signs we observed considerable divergence: signs that receive a low score in complexity have a poor performance in overall accuracy (e.g., SAUCE), and viceversa (e.g.,BONE) : see the tables below.

**Data-driven scale**

| | |
|---|---|
| FLOWER | 4,9 |
| RED | 4,7 |
| BAND_AID | 4,55 |
| PEAR | 4,55 |
| BONE | 4,5 |
| TREE | 4,4 |
| BREAD | 4,35 |
| GLASS | 4,2 |
| CASTLE | 4,1 |
| LEAF | 3,95 |
| PEN | 3,9 |
| FACTORY | 3,9 |
| THEATRE | 3,85 |
| SAUCE | 3,6 |
| COMPASS | 3,35 |

**Theory-driven scale**

| | |
|---|---|
| BAND_AID | 16 |
| SAUCE | 21 |
| TREE | 25 |
| CASTLE | 26 |
| RED | 26 |
| PEAR | 27 |
| GLASS | 28 |
| FLOWER | 29 |
| COMPASS | 32 |
| BREAD | 32 |
| THEATRE | 34 |
| LEAF | 35 |
| BONE | 35 |
| FACTORY | 36 |
| 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 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.

***References***: [1] Brentari, D. (1998). *A prosodic model of sign language phonology*. MIT Press. [2] Clements, G. N. (1985). The geometry of phonological features. *Phonology Yearbook* 2, 225-252. [3] Sagey, E. (1986). The representation of features and relations in nonlinear phonology. Doctoral dissertation, MIT, Cambridge, Mass.