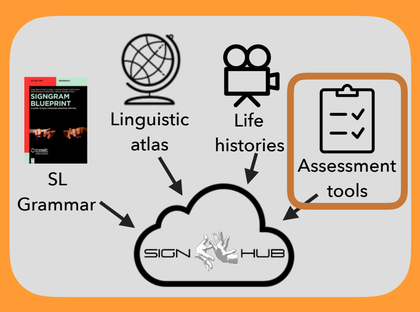
Measuring phonological complexity in sign languages

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

[1] LLF U. Paris, CNRS - [2] IJN ENS, CNRS - [3] U. Milano-Bicocca

The Sign-Hub project



Short-term / Scientific goal: assess the impact of delayed language exposure.

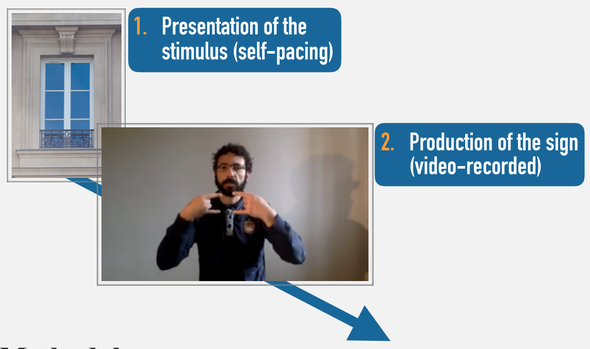
Long-term / Application: provide a basis for clinical assessment.

Battery of tests: lexical and syntactic tests.

LSF LIS LSC LSE ISL

Introduction

Assessing phonology: the production test in LSF and LIS.



Methodology:

- **Target signs:** non iconic, non transparent
- **Validation:** 20 hearing non-signers, exclusion if guessing >2
  - LSF: 112 signs, 4 excluded
  - LIS: 103 signs, 9 excluded

Setting the baselines:

- **Native** signers are a minority among signers
- Most Deaf people get to SL late or very late

➡ 3 Deaf populations per SL

15 natives | 15 early exposed (4yo-8yo) | 15 late exposed (9yo-13yo)

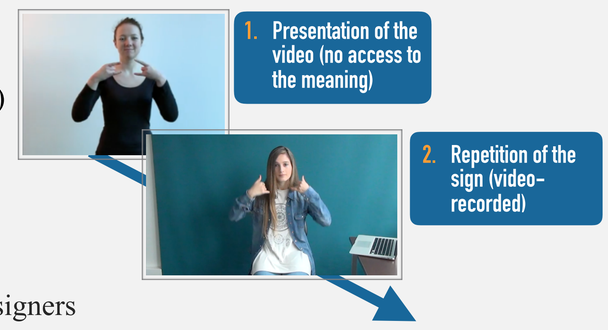
Factors:

- The frequency of each sign
- The **phonological complexity** of each sign, but:
  - No phonological description
  - No information on acquisition pattern or order
  - No recognition into speech/sign errors or other recognizable facts

Challenge: how to measure phonological complexity of signs without any knowledge of the phonology of the SL?

Non-linguistic measure

Repetition task for non-signers (video-recorded)



Materials:

- LSF:
  - 108 signs
  - 20 hearing non-signers
- LIS:
  - 94 signs
  - 17 hearing non-signers

Coding:

- 2 interns with basic competence in LSF/LIS + 2 SL researchers
- Fluency + Accuracy (Handshape (HS), Location (Loc), Orientation (Or), Movement (Mov))

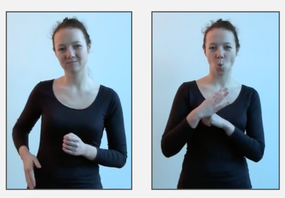
	Window
Fluency	1
HS	0
Loc	1
Or	1
Mov	1
TOTAL	4 / 5

Scoring:

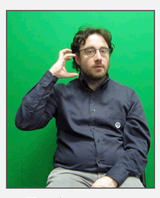
- Binary value for each component (correct = 1; wrong = 0)
- Overall accuracy: sum of the accuracy value for each component
- Degree of accuracy mapped onto a complexity scale (5 = least complex, 0 = most complex)

Results:

- Overall mean:
  - LSF: 4.282
  - LIS: 4.579


- Easiest: HAM (5)

- Most complex: HEDGEHOG (3.15)


- Easiest: RADIO (5)


- Most complex: HEDGEHOG (3.82)

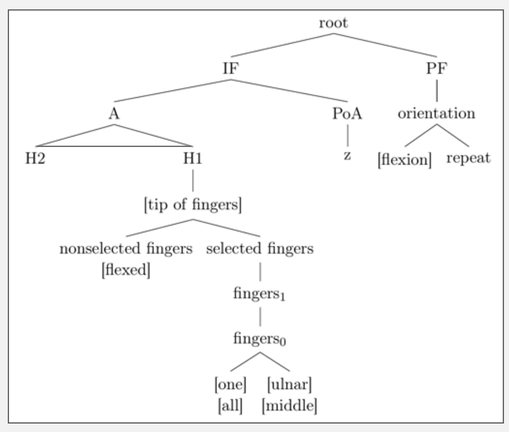
- Distribution of errors:
  - HS > Mov ≫ Or > Loc
- Mixed-model analysis:
  - Random effects: item & subject
  - Fixed effect: age
  - Main effect of age ($p < .05$)

Linguistic measure

Complexity scale: adapting the ***Prosodic Model*** (Brentari 1998)

Materials:

- LSF: 50 items annotated (/108), 3 excluded
- LIS: 30 items annotated (/94), 3 excluded



Coding:

- Tree structure for each sign

	Window
HS	16
Loc	2
Mov	4
TOTAL	22

Scoring:

- Level of complexity: number of nodes and positively specified features
- Lower values = less complex; higher values = more complex signs
- Total set of nodes and features: 116 (HS = 67, Loc = 22, Mov = 27)

Results:

- LSF


- Easiest: MATCH (14)


- Most complex: PEN (38)

- LIS


- Easiest: MOM (20)


- Most complex: PENCIL_SHARPENER (34)

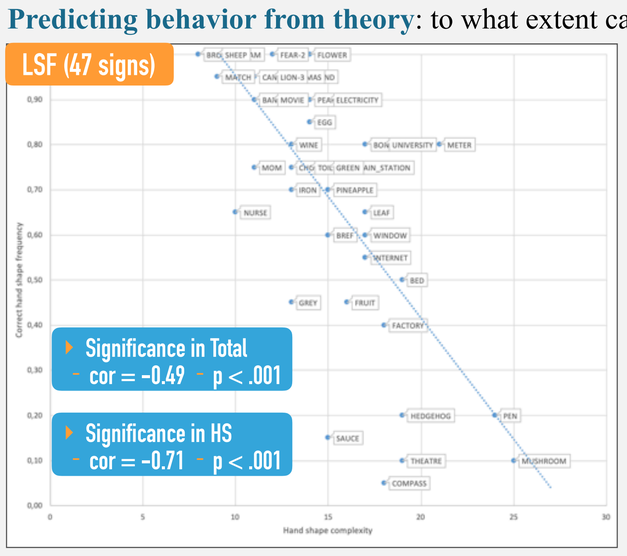
VIDEOS

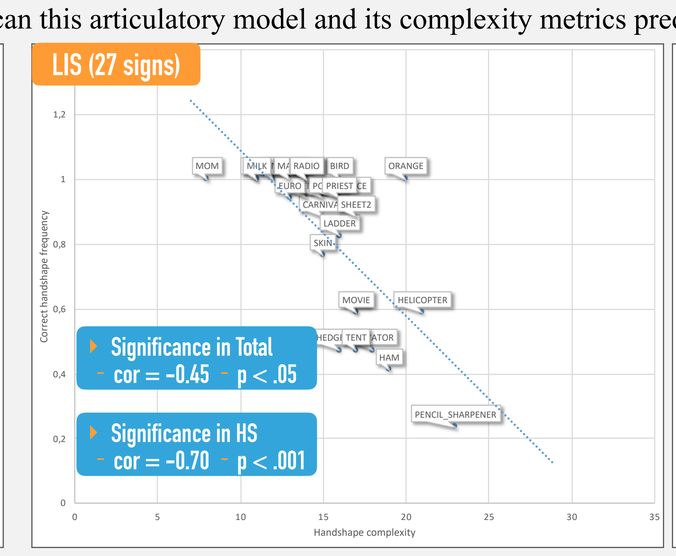
Correlation between data-driven & theory-driven complexity scores

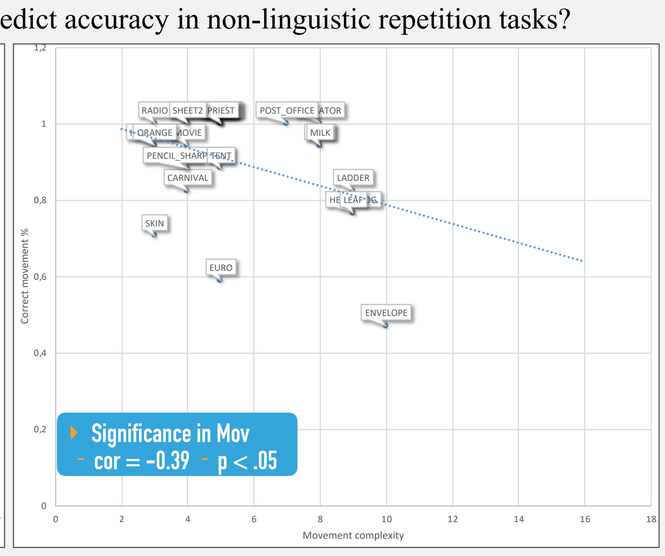
Predicting behavior from theory: to what extent can this articulatory model and its complexity metrics predict accuracy in non-linguistic repetition tasks?

A priori possibilities:

1. No relation (H0)
2. Full correlation for all phonemic classes
3. Correlation with some phonemic classes
4. Mixture of 2 and 3







Linear regression:

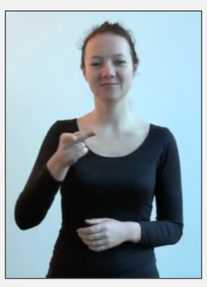
- No interaction

✓ HS: main effect

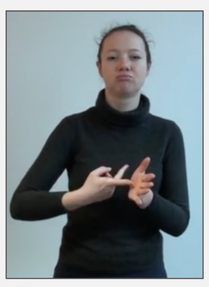
✳ Mov: main effect

Discussion

Some clear divergences in HS: cases where the predictions of the phonological model did not fit what was observed in the non-linguistic performance.


SAUCE

- Repeated **worse** than predicted by the model.
  - Data-driven HS: 0.15/1
  - Theory-driven HS: 15/25


COMPASS

- Repeated **worse** than predicted by the model.
  - Data-driven HS: 0.05/1
  - Theory-driven HS: 18/25


METER

- Repeated **better** than predicted by the model.
  - Data-driven HS: 0.8/1
  - Theory-driven HS: 21/25

Possible sources of mismatch:

1. Perceptual salience
2. Mismatch with gestural repertoire: [stacked]

1. Inhibition of movement
2. Different HS between the two hands
3. HS change not intuitive
4. Mismatch: [stacked]

1. No inhibition of movement
2. H2 open hand not complex
3. Activation of some iconic meaning

Conclusion

- LSF & LIS: HS is the most complex parameter and predicts complexity in a non-linguistic repetition task.
- LIS: Mov is also a factor of sign complexity.
  - ➡ Signs complexity might be due to only one parameter at a time.
- The non-linguistic measure is partially predicted by the phonological model.
  - ➡ This suggests a partial overlap between the phonology of signs and articulatory constraints applying on gestures.

Reference

Brentari, D. (1998). *A prosodic model of sign language phonology*. Mit Press.

Acknowledgments

Laurène Loctin, Deaf consultant

Valérie Guyot, Deaf consultant

Giuseppina Turco, PhD

European Commission | Horizon 2020 European Union funding for Research & Innovation

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 693349.

SIGN HUB

For further information about the project please visit our website at www.sign-hub.eu