

Автоматическое извлечение фонемного инвентаря



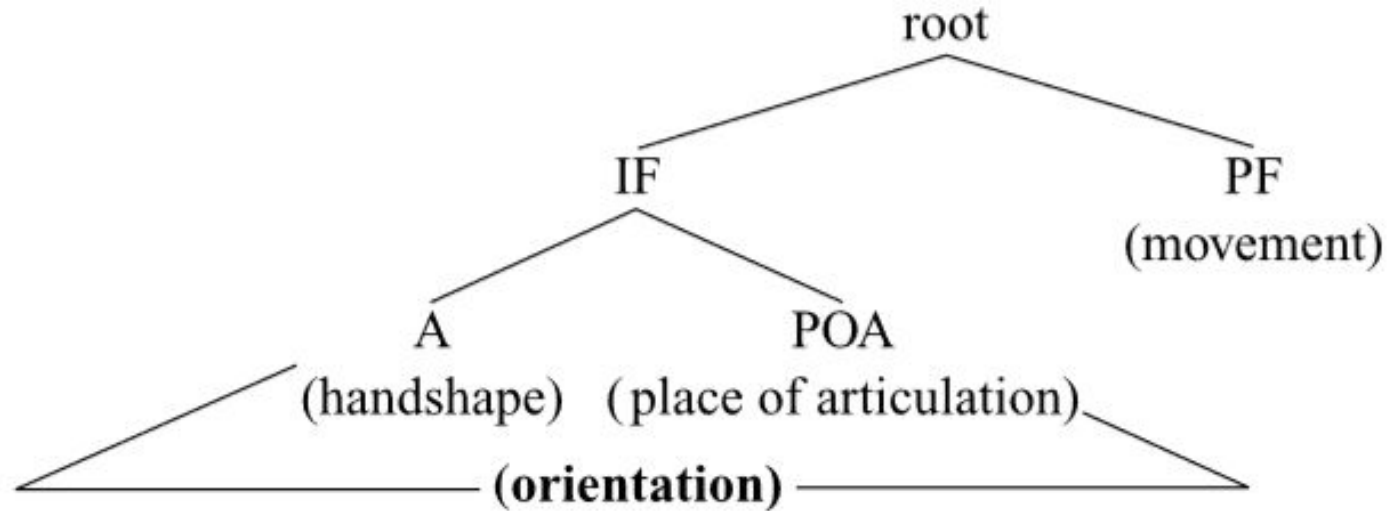
15.02.2019 Новосибирск
Анна Клезович (НИУ ВШЭ)

Развитие фонологии жестовых языков

- 1) Stokoe
- 2) Movement-Hold theory Liddell&Johnson
- 3) Hand-Tier theory Sandler
- 4) Dependencies Model van der Hulst
- 5) **Prosodic model of phonology Brentari**

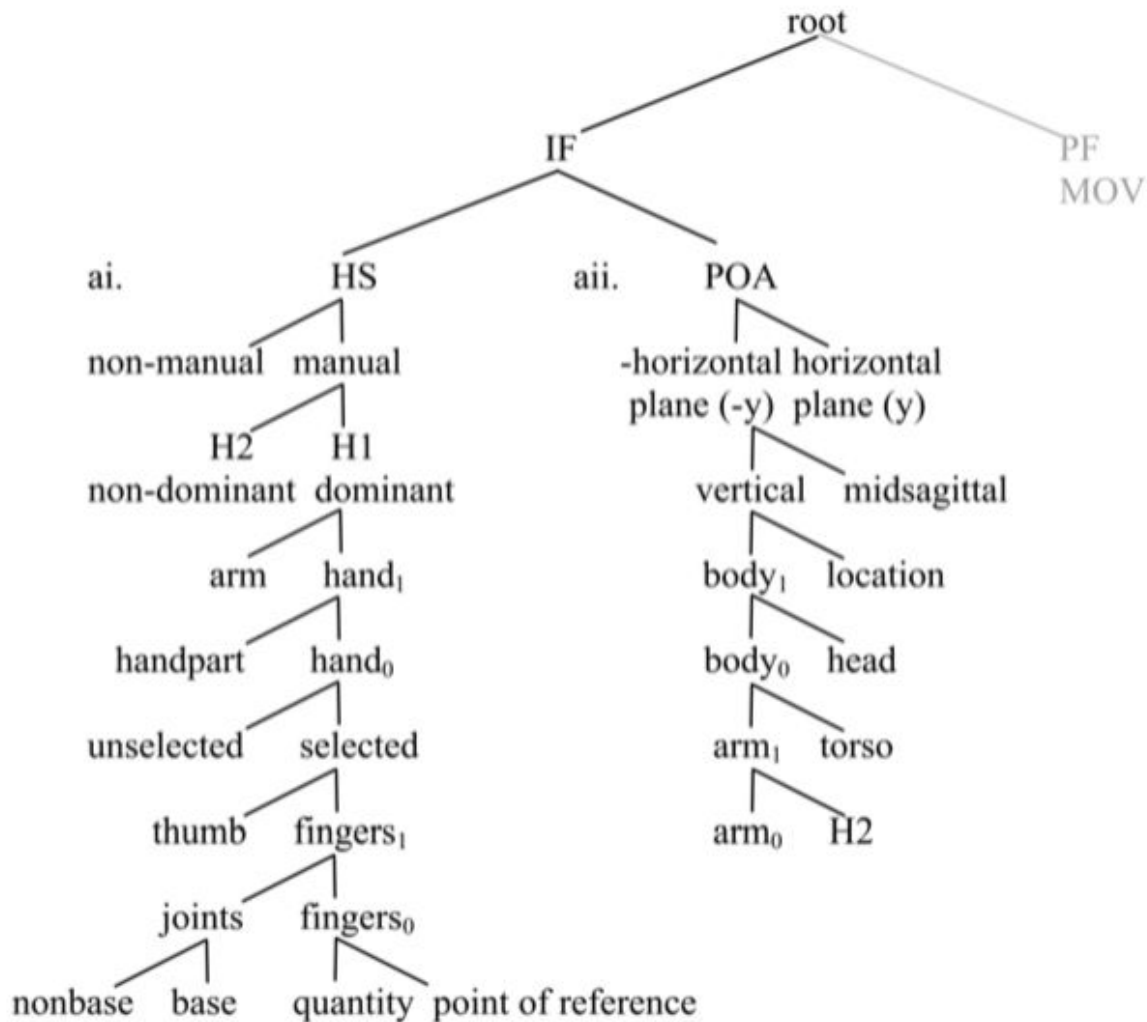
Prosodic Model

b. *Parameters in the model*

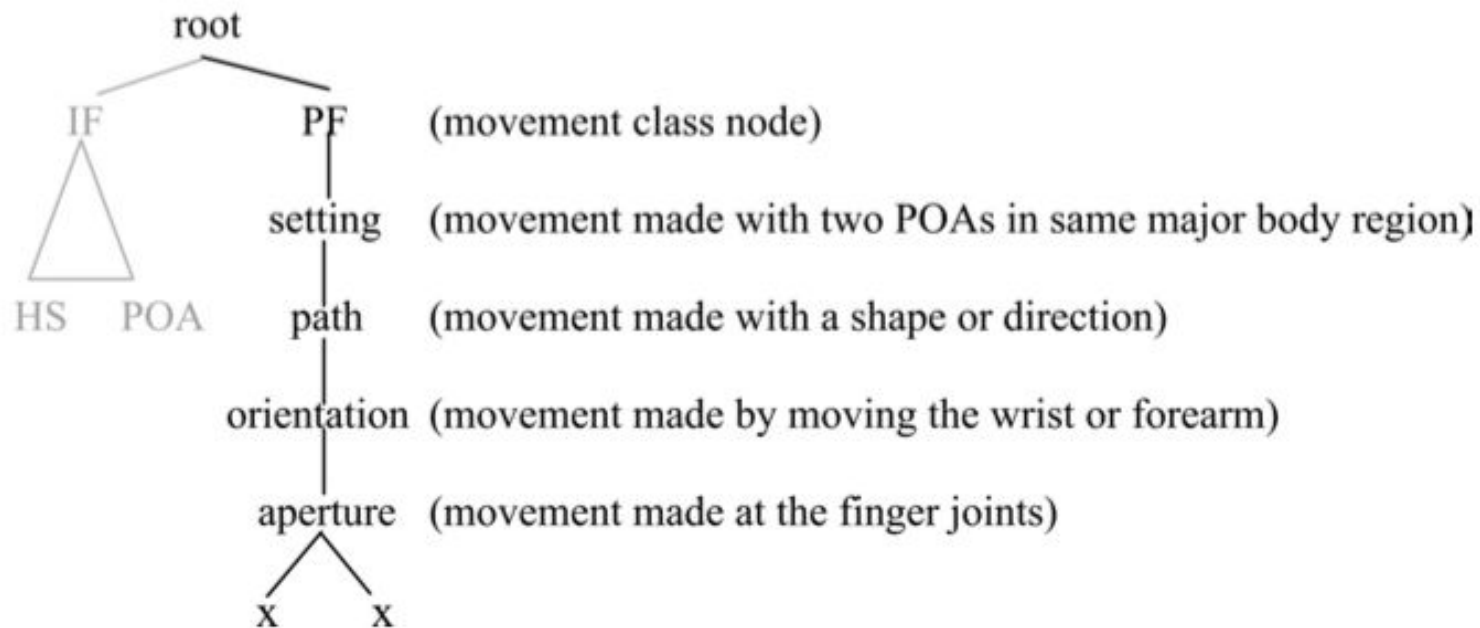


- Inherent features - something that cannot change
- Prosodic features - something that does change

Inherent features

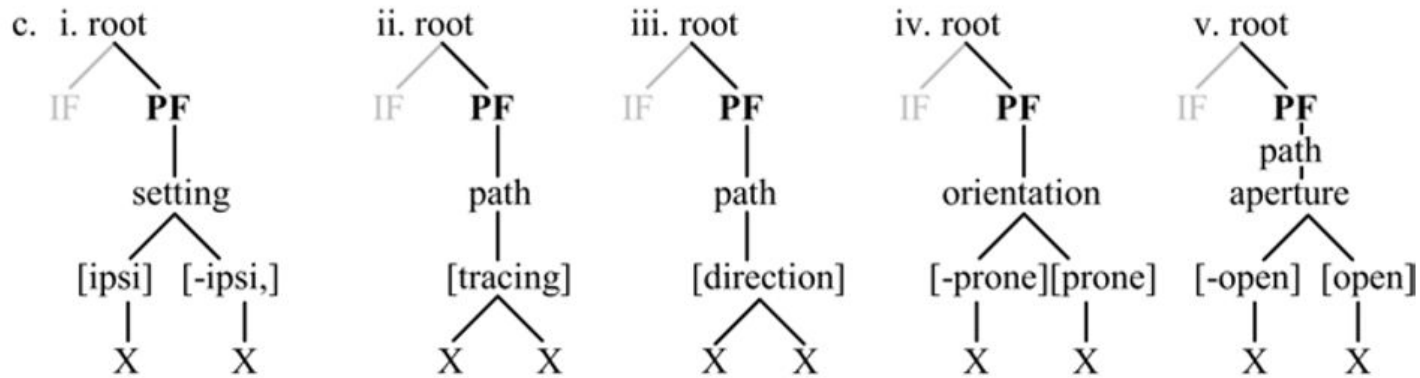
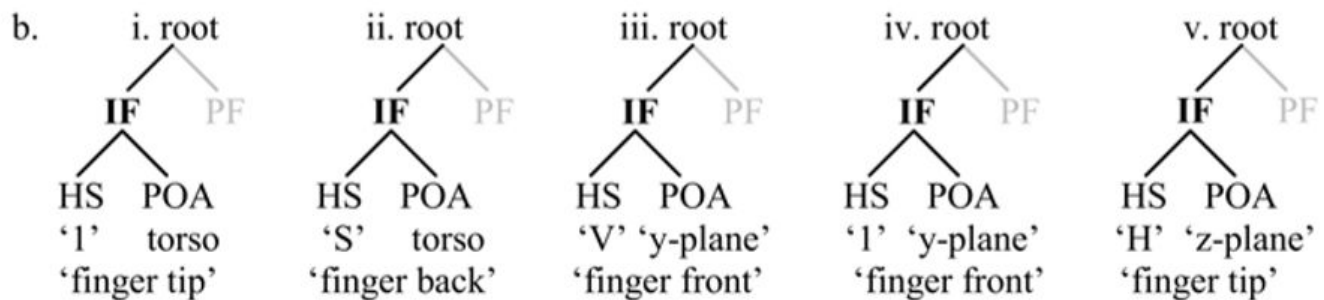


Prosodic features





a. i. WE ii. SORRY iii. SIT iv. HAPPEN v. THROW



Основная идея

Автоматически извлекать позиции без движения (холды) для большого количества жестов. Просмотреть получившиеся холды и отобрать одинаковые. Это и есть фонемный инвентарь.

Как извлекать?

Calle Börstell's code



https://github.com/borstell/make_sign_stills

Мой апдейт на **Calle Börstell**

- Для своего датасета изменить hardcoded figures опытным путем
- Понять, можно ли их автоматически подстраивать для любого датасета, иначе говоря для любого языка
- Есть ли зависимость от языка?

Open-CV versus openPose

https://docs.opencv.org/master/d6/d00/tutorial_py_root.html

<https://github.com/CMU-Perceptual-Computing-Lab/openpose>

Методология

1. Automatically retrieve list of words from spreadthesign online dictionary (done)
2. View all signs and manually delete compounds, phrases, numbers, dactyl, signs based on dactyl letters. (1915 done. There seems to be about 5000 thousands signs which suit purposes of this research)
3. Disyllabic signs need to be treated differently. Maybe also deleted from the dataset
4. Download automatically signs by meanings which were left after manual cleaning
5. Using API for spreadthesign.

Методология

6. Adjust make-signs-still.py program for Russian Sign Language and my dataset
7. With the help of the program retrieve all hold positions
8. Manually go through all retrieved hold positions and group them into categories with the same orientation and hand configuration. These have to be analyzed as phonemes.
9. Also try openPose module for python in case there is more powerful computer or gpu on online clouds will turn out to be enough. This is important to do in order to catch complex movements which cannot be derived just from the knowledge of starting and ending hold positions.

Проблемы методологии

- Фразы, компаунды, числительные, все производное от числительных, дактиль удаляется вручную из датасета (это долго!)
- Почему удаляю именно это?
- Но если не удалять, то обрабатывать автоматически такие жесты еще дольше!

Проблемы методологии

2.2. Data Processing

Video files were downloaded from the *Spread the Sign* public website, along with metadata on the language used, as well as the name of the concept and the concept's category (in English)⁵. As the first step in our process, we used the model of [Cao et al. \(2016\)](#) to identify the position of video frames. The model file published by the authors, which has been used in our work, is highly accurate in identifying the required body parts. This step required about **two months** of computing time on a single GPU. 

<https://www.frontiersin.org/articles/10.3389/fpsyg.2018.00725/full#B31>

Проблемы методологии

- Disyllabic signs (Пример) Решение?
- Triled movement (TM) **НО**: epenthetic movement (Пример) Решение?
- Complex movement (Пример) Решение?

Следствия для теории

Established phonemic inventory of RSL

How well each of the previous theories applies to RSL.

Maybe I will have some insights into how to develop the Prosodic Model theory.

Следствия для практики

If we know phonemes and phonological processes which can apply to these phonemes we can analyze all signs as a list of phonemes arranged in some order, so that we can recognize not whole signs but phonemes on the videos. In addition to that, the phonemes inventory has to be obviously much shorter than the number of all signs in RSL. Thus, we will be able to use less computer memory and make the process of signs recognition more effective.

References

- Brentari, Diane (1998) The Prosodic Model of Sign Language Phonology. MIT press.
- Börstell, Calle (2018). make-signs-still.py.
https://github.com/borstell/make_sign_stills
- Hulst, Harry van der (1995) The Composition of Handshapes. In: University of Trondheim Working Papers in Linguistics 23, 1–18.
- Liddell, Scott & Robert E. Johnson (1994) American Sign Language: The Phonological Base. In: Sign Language Studies 64, p. 195–278.
- Sandler, Wendy (1989) Phonological Representation of the Sign: Linearity and Nonlinearity in American Sign Language. Dordrecht: Foris Publications.
- Stokoe, William (1960) Sign Language Structure: An Outline of the Visual Communication Systems of the American Deaf. In: Studies in Linguistics, Occasional Papers 8. Silver Spring, MD: Linstok Press.