尤洁雪
Beijing Language and Culture University, 2017/2018

波兰语的复辅音 - 分析与可视化

Analysis and visualisation of consonant clusters in Polish language based on National Corpus of Polish

Background

In linguistics, a consonant cluster, consonant sequence or consonant compound is a group of consonants which have no intervening vowel. In English, for example, the groups /spl/ and /ts/ are consonant clusters in the word splits.

Many Slavic languages may manifest almost as formidable numbers of consecutive consonants, such as in the Slovak words štvrť /ʃtvr̩tʲ/ ("quarter"), and žblnknutie /ʒbl̩ŋknutje/ ("clunk"; "flop") and the Slovene word skrbstvo /skrbstʋo/ ("welfare"). However, the liquid consonants /r/ and /l/ can form syllable nuclei in West and South Slavic languages and behave phonologically as vowels in this case. An example of a true initial cluster is the Polish word wszczniesz (/fʂt͡ʂɲɛʂ/ ("you will initiate").

A phonological word may begin with a five-consonant cluster: /"spstroNĺem/ z pstragiem 'with (the) trout', and a 4C lexeme-initial cluster is not unusual: /vzglont/ wzglad 'respect (n.)'. Even a geminate affricate can occur in this position: /"SSkstk/ dzˑdzˑysty 'rainy'. According to W. Jassem the initial consonant clusters are not simplified in fluent, spontaneous speech.[1]

Research topic:

In this project I will analyse National Corpus of Polish in terms of consonant and consonant cluster distribution. Based on the corpus data I will try to find most common cluster models and their placement in a prosodic word (initial, mid-word, final).

I will test the hypothesis that prosodic words containing C-clusters are avoided in speech as opposed to written Polish due to minimising articulatory effort. For testing this hypothesis I will compare consonant cluster distribution in written and spoken Polish based on corpus data.

I plan to measure, analyse and visualise following data:

- vowel to consonant ratio 元音辅音对比
- common consonant clusters (rated by frequency) 最常见的复辅音
- consonant cluster models (rated by frequency) 复辅音模型
- consonant cluster placement (initial, mid-word, final) 复辅音地位 （词头，词中，词尾）
- distribution of consonant clusters in written and spoken Polish 波兰语口语与书面语的复辅音分布，统计分析

Material:

http://www.nkjp.uni.lodz.pl/index.jsp

References:

1. Piotr Pęzik (2012) Wyszukiwarka PELCRA dla danych NKJP. Narodowy Korpus Języka Polskiego (Przepiórkowski A., Bańko M., Górski R., Lewandowska-Tomaszczyk B (red.). 2012. Wydawnictwo PWN . Do pobrania.
2. Eugeniusz Cyran and Edmund Gussmann, Consonant clusters and governing relations: Polish initial consonant sequences.
3. W. Jassem , Illustrations of the IPA: Polish, Institute of Fundamental Technological Research Polish Academy of Sciences