How Predictability Shapes the Way We Speak: Lessons from Polish Conversations and Reading Aloud

This study investigates how contextual predictability—quantified as phoneme-level surprisal—influences acoustic realization in Polish read and conversational speech. While predictability-driven variation in speech is well-documented in languages with acoustically prominent stress systems, such as English, its manifestation in languages like Polish—with fixed penultimate stress and relatively weak acoustic stress marking—remains understudied. Polish offers a valuable opportunity to explore how prosodic expectations interact with information-theoretic pressures in speech production. We ask whether surprisal affects segmental duration and vowel space dispersion in Polish, and how these effects interact with lexical stress, sentence accent, and discourse type.

We analyze data from the PRODIS corpus, a large-scale, phonemically aligned Polish speech resource containing recordings from 50 native speakers. The dataset comprises read speech from four Wikipedia domains (science, politics, history, culture) and spontaneous conversations with an experimenter. To estimate contextual predictability, we use two phoneme-level language models based on a lightweight GPT-2 architecture (*NanoGPT*; Karpathy 2025), trained separately on phonemized Polish Wikipedia and OpenSubtitles corpora. This dual-model design captures discourse-specific structures in written and conversational language (Foremski in prep).

Our analysis targets two key acoustic measures: segmental duration and vowel formant dispersion (as an index of vowel space expansion). In read speech, we find that phonemes with higher surprisal are produced with longer durations and more peripheral vowel realizations. These patterns are consistent with the Smooth Signal Redundancy (SSR) hypothesis, which posits that speakers increase acoustic effort for less predictable segments to maintain intelligibility.

Importantly, these effects persist despite the weak surface marking of stress in Polish. Stressed syllables show greater acoustic enhancement overall, which is further amplified when surprisal is high. This suggests an interaction between prosodic structure and contextual predictability in shaping articulation. Lexical frequency also has a small but independent effect, with more frequent words showing shorter durations and more centralized vowels.

Surprisal effects vary across discourse topics. Scientific texts exhibit stronger acoustic enhancement of high-surprisal segments than historical or cultural texts, suggesting genre-specific adaptations in speech planning, possibly linked to differing informational demands or audience expectations.

Analysis of the conversational speech data is currently underway. Preliminary inspection suggests similar, though potentially attenuated, effects of surprisal on acoustic distinctiveness in accordance with van Son and Pols (2004), and Ernestus (2014). This points to genre-dependent speech planning strategies, possibly reflecting variation in information density or listener expectations across domains.

In sum, this study provides robust evidence that contextual predictability significantly influences segmental duration and vowel articulation in Polish, despite its relatively weak acoustic stress system. These findings support cross-linguistic models of prosodic-acoustic adaptation and highlight the importance of integrating discourse-level and probabilistic

factors into models of speech production. The work contributes valuable empirical data to the growing body of research on acoustic correlates of information structure in underexamined language families such as the Slavic group.

References:

Ernestus, M. 2014. “Acoustic reduction and the roles of abstractions and exemplars in speech processing’”. *Lingua*, *142*, 27-41.

Foremski, J. In preparation. “Contextual predictability and speech variation: A study of read and spontaneous Polish using discourse-aware language models” [Przewidywalność kontekstowa a zmienność mowy: Badanie polszczyzny czytanej i spontanicznej z użyciem modeli językowych uwzględniających dyskurs]. Master thesis to be defended at Faculty of English, Adam Mickiewicz University, Poznań, Poland.

Karpathy, A. 2025. “Nano-GPT,” https://github.com/karpathy/nanoGPT

Van Son, R., Bolotova, O., Pols, L. C., & Lennes, M. 2004. “Frequency effects on vowel reduction in three typologically different languages (Dutch, Finish, Russian)”. In *Interspeech* (pp. 1277-1280).