Articulatory and acoustic variation in the realization of Polish retroflexes

Anita Lorenc[1], Marzena Żygis[2], Łukasz Mik[3], Daniel Pape[4], Márton Sóskuthy
[1]University of Warsaw, [2]Leibniz-Centre General Linguistics, Berlin, [3]State Higher Vocational School, Tarnów; [4]McMaster University, Hamilton, [5] The University of British Columbia, Vancouver

Traditionally, Polish postalveolar sibilants were described as palatoalveolars /ʃ, tʃ, ʒ, dʒ/ (Dukiewicz 1995). However, recent research has revealed that the sounds are rather retroflexes from both acoustic (Żygis, Pape & Jesus 2012) and articulatory point of view (Bukmaier & Harrington 2016), concluding that Polish sibilants are produced with tongue tip posture and can be characterized as retroflexes based on the tongue tip orientation.

Since the conclusion was mainly drawn on the tongue tip characteristics (Bukmaier & Harrington 2016) we aim to gather more articulatory evidence regarding the *whole* tongue shape and additional articulators, i.e., jaw, lips. The second aim of the current study is to further examine the relationship between articulation and acoustics, based on previous findings (see Pape & Żygis 2016) according to which acoustic spectra of Polish retroflex voiceless sibilants show an additional major spectral peaks around 7kHz that cannot be explained by front or back cavity resonances. However, in a previous articulatory synthesis modelling study (Pape & Żygis 2016 based on Birkholz 2014) it was shown that horizontal lip protrusion, but not lip spreading, induced a second major peak in the acoustic spectrum around 7kHz. With the present dataset we aim to verify whether the articulatory synthesis modelling would match real world articulatory lip data.

In order to examine the outlined research questions we conducted a study with 20 rigorously selected speakers of Standard Polish. The data were obtained by means of articulography (Carstens AG500) combined with acoustic recordings and three Point Grey cameras. The EMA sensors were placed on the tongue (tongue tip (TT), tongue front (TF), tongue dorsum (TD), tongue back (TB), lips (upper lip (UL), lower lip (LL)), and jaw (J), and three reference sensors for movement correction (see fig.1). Our material consisted of three syllabic words with retroflexes ([d͡ʐ, ʐ, t͡ʂ, ʂ]) appearing intervocalically in [a] context in word-medial stressed position, e.g. *kaszalot* [kaʂalɔt] "cachalot". For comparative analysis, words with the voiceless dental plosive in the same context were also used, e.g. *latarka* [latarka] "torch".

Our results reveal that all speakers (with one exception), show a higher position of TT in the production of retroflex consonants compared to [t] (e.g. mean difference tongue tip elevation is 2.73 mm for [ʂ] compared to [t]). Furthermore, all speakers retract the back of the tongue (TB) in the production of retroflex sibilants compared to [t] (mean: 7.79 mm). Finally, all speakers display greater distance between TT and jaw in the case of retroflex sibilants compared to [t] (e.g. 15.69 mm for [ʂ], see fig.2). The obtained results reveal both inter- and intraspeaker variability in the place of articulation of Polish retroflex sibilants being either alveolar or postalveolar. There were also differences in the tongue shape ranging from completely flat to slightly concave (e.g. [ʂ] mean 0.98 mm) to slightly convex (e.g. [ʂ] mean: 0.78 mm). Speakers also show different degrees of jaw lowering.

To answer the second research question, we used the *relative* UL and LL lip position (horizontal, vertical, Euclidean) in order to measure the absolute lip gesture amplitude for each speaker with respect to a 'neutral lip position'. These relative lip values were then correlated with the magnitude of the second major spectral peak in the acoustic spectra at various timepoints (see fig. 3). Our results show that (1) very strong inter-speaker and intra-speaker variation in lip movement during sibilant production is found, (2) all correlations between articulation (relative lip movement) and acoustics (second spectral peak) are rather weak and not significant and that (3) variation is not substantial comparing the spectra at the acoustic midpoint with the 25% and 75% sibilant duration timepoint.

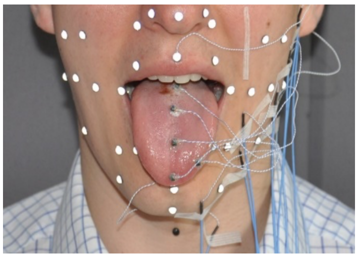

Figure 1. Articulatory and camera sensor placement for the recording setup.

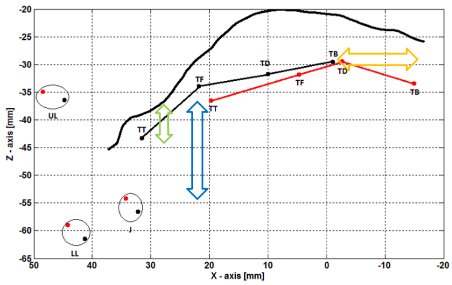

Figure 2. Palate tracing and EMA sensors positioning at the moment of minimum velocity (MinVEL) reached by the tongue sensors (TT, TF, TD, TB) in the nuclear phase of consonant [ʂ] (red line) and [t] (black line) for speaker F10 and differences in sensor position (TT [ʂ] – TT [t] green arrow, TT [ʂ] – J [ʂ] blue arrow, TB [ʂ] – TB [t] yellow arrow.

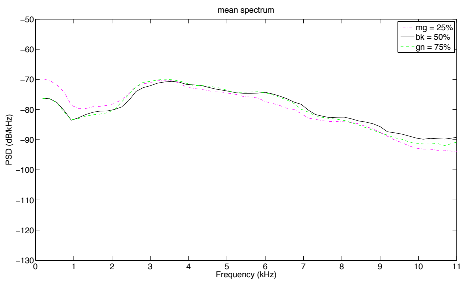

Figure 3: Mean acoustic spectrum (all speakers) at the acoustic midpoint (black line), the acoustic 25% point of sibilant duration (magenta line) and the 75% (green line) of the sibilant [ʂ] (target word kaszaki [kaʂaki]).

References:

Birkholz, P. 2014. Enhanced area functions for noise source modeling in the vocal tract. *Proceedings of the 10th International Seminar on Speech Production*, Cologne.

Bukmaier, V., & Harrington, J. 2016. The articulatory and acoustic characteristics of Polish sibilants and their consequences for diachronic change. *JIPA*, 46(3), 311–329.

Dukiewicz, L. 1995. Fonetyka. In: Dukiewicz, L. & I. Sawicka. Fonetyka i fonologia.Warszawa: Instytut języka polskiego.

Hamann, S. 2003. The Phonetics and Phonology of Retroflexes. Utrecht: LOT.

Pape, D., & Żygis M. 2016. The puzzle of two major spectral peaks in Polish sibilants: Acoustic Analysis and Articulatory Synthesis. Poster presented at the Satellite Workshop on "Dynamics and Representation of Turbulent Sounds. Laboratory Phonology Conference 15, Cornell University.

Żygis, M., Pape, D. & L. Jesus. 2012. (Non)retroflex Slavic affricates and their motivation. Evidence from Czech and Polish. *JIPA 42*. 281-329.