The puzzle of two major spectral peaks in Polish sibilants: Acoustic analysis and articulatory synthesis

Daniel Pape (McMaster University, Hamilton, Canada),
Marzena Żygis (ZAS & Humboldt University, Berlin, Germany)

Background

- Sibilants with postalveolar place of articulation are characterized by one broad spectral peak in the mid-frequency region (3-7kHz)
- Main peak is attributed to front cavity resonance (Stevens 1998, Shadle et al. 2009) whereas back cavity is assumed to be acoustically inactive (Stevens 1998)
- Centre frequency of the prominent spectral peak codes how fronted sibilant is produced -> more fronted PoA resulting in higher frequencies
- Original research question: How are the spectral characteristics of sibilants affected by both prosodic contrast and place of articulation (see Żygis, Pape, Jesus & Jaskula 2014, 2015)?

Production experiment

- Speech production experiment:
  - 16 **Polish** speakers (8 males, 8435 items)
  - Polish **retroflex and alveolo-palatal fricatives** in word-final and sentence-final position
  - Monosyllabic target words embedded in frame sentences
  - Two **prosodic conditions**:
    - polar questions ("Widzi ten pła[ʂ͡tʂ]?")
    - statements ("Widzi ten pła[ʂ͡tʂ].")
- **Results:**
  - One major peak (corresponding to front cavity resonance) around 3kHz
  - **But additionally** an unexpected major spectral peak around 7kHz
  - High energy for this peak, often more energy than in the first peak

Comparison: German speech productions

- **Question:** Is this additional spectral peak found in other languages as well?
  - Pilot study with German postalveolar fricatives (5 speakers, 164 items)
  - Result: No trace of additional peak

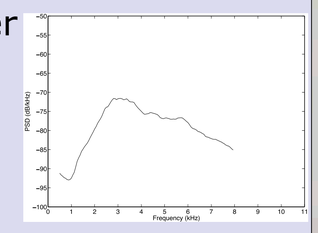

Figure 1: Mean multitaper spectrum of real speech productions of German postalveolar fricative /ʃ/

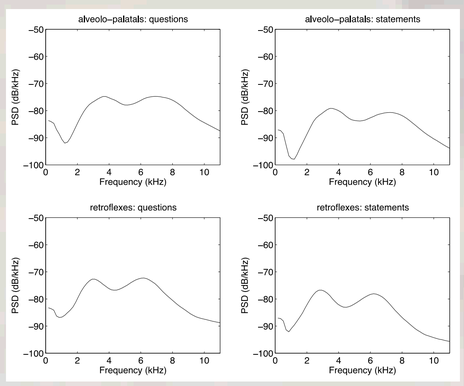

Figure 2: Mean spectra of Polish alveolo-palatals (upper panels) and retroflexes (lower panels), split by prosodic condition.

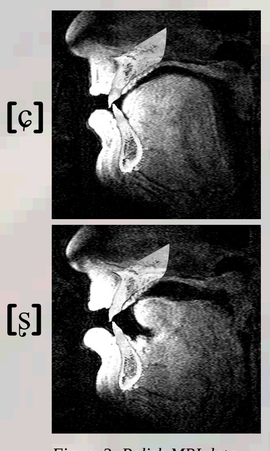

Figure 3: Polish MRI data recorded by Martine Toda @ATR Japan; with permission

Reasons for the second peak

We pose two possible explanations for the second major peak:
1. **Front cavity resonances**: second peak is higher resonance of the dominant front cavity
2. **Other cavities:** second peak is generated by resonances or additional modes in other cavities (lip-teeth, back)

Hypothesis 1: Front cavity

Correlations between first and second major peak:

- If second peak is a higher resonance of the dominant front cavity then high correlations between first and second peak should occur (since second peak is a multiple of first peak)
- We computed heat maps (two-dimensional map where the number of pair occurrences is coded by color intensity)
- **Results:** no apparent linear relationship between first and second peak, weak correlation coefficient (r=-0.26)

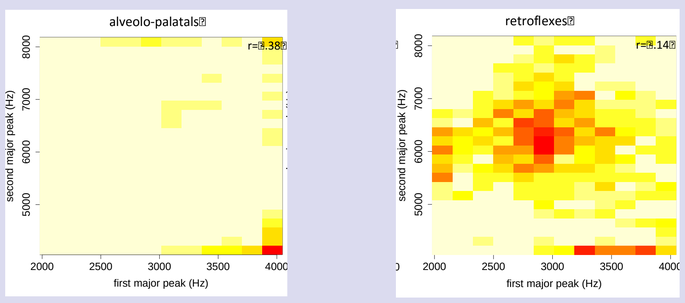

Figure 4: Heat maps for peaks of Polish alveolo-palatals and retroflexes: pairs of the first major spectral peak (x-axis) and the second major spectral peak (y-axis). The correlation coefficients between the two dimensions are printed on top of the corresponding figure. For each given x-y pair, darker colors mean higher item counts (i.e., more occurrences of this specific pair).

Hypothesis 2: Other cavities

Lip-teeth cavity

- Simulations with articulatory synthesizer VocalTractLab 2.1 (Birkholz 2014)
- Real MRI vocal tracts of standard German postalveolar sibilant
- Stepwise manipulation of lip protrusion and lip opening
- Results:
  - **Increasing lip protrusion generates a strong additional peak at 5-6kHz**
  - Decreasing lip protrusion and/or lip opening had no effect

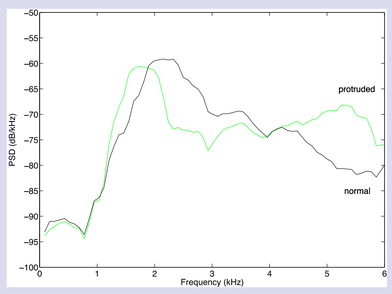

Figure 5: Multitaper spectra of synthesized simulations (Vocaltractlab) with normal (black) and protruded (green) lips.

Back cavity

- Stepwise manipulation of back cavity size
- Results:
  - The manipulation of the **back cavity size** (from very small to very large) had **no visible effect on higher frequencies**

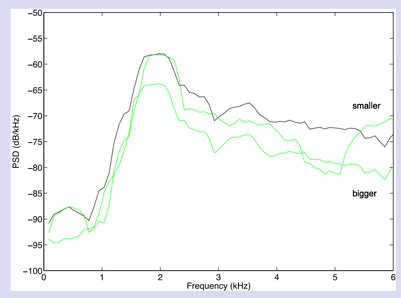

Figure 6: Multitaper spectra of synthesized simulations (Vocaltractlab) of the standard back cavity size (black) in comparison to decreased and increased back cavity size (green).

CONCLUSIONS

1. The frequency of the second peak is independent from the frequency of the first peak. We conclude that the **front cavity hypothesis should be rejected**.
2. The **traditional approach of using centre of gravity (CoG)** to code the PoA (i.e. ignoring the second major peak) **is highly problematic** for Polish fricatives and should be used with caution.
3. Our results show that the **second major peak in Polish retroflexes and alveolo-palatals** can be **attributed to** a **strong articulatory lip protrusion**.
4. The simulations of back cavity manipulations show no effect on higher frequencies. We therefore conclude that this part of the hypothesis should be rejected.

References

Birkholz, P. (2014). "Enhanced area functions for noise source modeling in the vocal tract". Proceedings of the 10th ISSP, Cologne.
Shadle, C., Proctor, M., Berezina M. & Iskarous, K. (2009). "Revisiting the role of the sublingual cavity in the /s/ - /ʃ/ distinction", 157th Meeting of the ASA, Portland, USA.
Stevens, K. N. (1998). Acoustic phonetics, MIT Cambridge.
Toda, M., Shinji, M., & Kiyoshi, H. (2005). "Formant-cavity affiliation in sibilant fricatives". In Fuchs, S., Toda, M. & Zygis, M. (eds.), Turbulent sounds: An interdisciplinary guide. Berlin: de Gruyter. 343-374.
Żygis, M., Pape, D. & Jesus L.M.T. (2012). "(Non)retroflex Slavic affricates and their Motivation. Evidence from Czech and Polish", *Journal of the International Phonetic Association* 42(3), 1-49.
Żygis, Marzena, Pape, Daniel, Luis M.T. Jesus & Marek Jaskula (2014). " Intended intonation of statement and polar questions in Polish whispered, semi-whipered and normal speech modes". *Proceedings of the 7th International Conference on Speech Prosody*, Dublin. Irland. 678-682.
Żygis, Marzena, Pape Daniel, Luis Jesus & Marek Jaskula (2015). "The effects of intonation on acoustic properties of fricatives". *Proceedings of the 18th International Congress of Phonetic Sciences*. Glasgow, UK. Paper number 481, 1-4.

This research has been supported by Bundesministerium für Bildung und Forschung (BMBF, Germany) Grant Nr. 01UG1411 to Marzena Żygis.