Background: subglottal resonances

- Resonances of the human subglottal tract are fairly constant for a given speaker (no moving articulator) [1, 2].
- Subglottal resonances (SGRs) can distort spectral peaks of formants [1], therefore speakers avoid putting vowel formants in these regions [2].
- SGRs have been claimed to be natural divisions between +/− value of several distinctive features, e.g.:
  - Sg1 is a boundary between low and non-low vowels.
  - Sg2 is a boundary between back and non-back vowels.
  - American English [2], Spanish [4], German [5], Korean [6], Hungarian [7, 8].
- However, some vowel realizations contradict the subglottal hypothesis: Hungarian [7, 8]:
  - Some speakers produced some/most of their /ow/ [i] & [o] with F1 < Sg1.
  - back [a] with F2 > Sg2 and front [a] & [e] with F2 < Sg2.
- Hypothesis: Back vowel-context dependency vowel target underhe and consequent coarticulation is responsible for moving the formant-SGR relations.

Methods: recordings & measurements

- 2 male (m1, m2) and 2 female (f1, f2) adult native speakers of Standard Hungarian.
- Utterances “CVC’P” nonsense words in the career sentence: “Majtot a Ccqv/ov staying over...” [1] is the reading the word C.CV/’P none.
- 4x4x4x4 consonant contexts.
- 8 male, 8 female, 8 low/mid-low, 8 back, rounded, [o] unrounded, [back], but phonetically front for most speakers [9, 10].
- 8x10, mid-low, front, rounded.
- 6 repetitions per nonsense word.
- Simultaneous microphone and accelerometer recordings.
- First and second formant (F1, F2) from microphone signals:
  - Measured semi-automatically using Praat + manual correction.
  - Measured 24 times during each vowel at regular intervals.
  - Measured at time of highest F1 (Fig. 2 - see Fig. 2).
- First and second SGRs (Sg1, Sg2) from accelerometer signals:
  - Measured manually 25 times for each speaker.
  - Means were considered ground truth.

Results

1. Inter-speaker effects:

![Figure 1: Vowel spaces for F1 vs. F2. The vertical dashed lines indicate the speakers’ mean Sg2 +/− one standard deviation. The horizontal dashed lines indicate the mean Sg1 +/− one standard deviation.](image)

- F1 on the expected side of Sg1:
  - [i] 0.94−9.5%,
  - [a] 79.1% for m1, 100% for the other 3 speakers,
  - [e] 0−9.8%.
- F1 on the expected side of Sg2:
  - [a] 76.1% for m1, 100% for the other 3 speakers.
- F2 on the expected side of Sg2:
  - [a] 100% for all speakers,
  - [e] 90.8% for m2.
- F2 of [a] & [e] show no recalcitrance in the present study.

2. Context effects:

![Figure 2: Example average formant trajectories based on the results for speaker f2 - representative for our present subjects.](image)

- (j) & (g) pull the F1 of [a] closer to /o/ on the wrong side of the SGR.
- (j) & (g) pull the F2 of [a] closer to /o/ on the wrong side of the SGR.

3. Both effects are enhanced if both context (m) and/or (j).

4. Chi-square test: significant effect of context across all speakers:

![Figure 3: Results of the χ2 analysis.](image)

- m1 vs. m2 f1 < 0.05 10.885 0.143 12.518 0.157
- F1 of [a] < 0.03 9.000 0.152 13.241 0.184
- F2 of [a] < 0.001 114.185 0.548 41.165 0.327

Conclusions & further questions

- The recalcitrant vowels [i] & [o] showed both speaker and context dependency:
  - The data suggest that coarticulatory context effects can mask the formant-SGR relations. However, more data from more speakers are needed to be able to draw general conclusions.

- The vowels were on the wrong side of the SGRs less often than in previous studies:
  - This could be due to inter-speaker differences or to differences in phonetic context (unstressed syllables in the former studies, stressed syllables in the present study).

- What impact do these context- and speaker-dependent effects have on vowel perception?

Acknowledgements

T. E. W. was supported by ISCA- Google grant. M. L. was supported in part by NSF Grant No. 0950210. T. G. C. was supported by the Alcatel-Alcatel-Lucent 2005-2006, the TAMOP-4.2.2.1-11/1008-2007 project, and by the TAMOP-4.2.1-11/1008-2007 project.

Key references