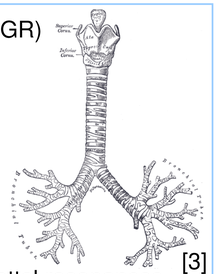
Relation of formants and subglottal resonances in Hungarian vowels

Tamás Gábor Csapó[1], Zsuzsanna Bárkányi[2], Tekla Etelka Gráczi[2], Tamás Bőhm[1,3], Steven M. Lulich[4]

[1]Department of Telecommunications and Media Informatics, Budapest University of Technology and Economics, Hungary
[2]Research Institute for Linguistics, Hungarian Academy of Sciences, Budapest, Hungary
[3]Institute for Psychology, Hungarian Academy of Sciences, Budapest, Hungary
[4]Speech Communication Group, MIT, Cambridge, MA 02139

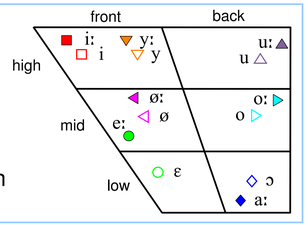
csapot@tmit.bme.hu, bzsu@budling.nytud.hu, graczi@nytud.hu, bohm@tmit.bme.hu, lulich@speech.mit.edu

Background: subglottal resonances (SGR)


[3]

- resonances of the human subglottal tract, [1] (Sg1~600Hz, Sg2~1400Hz and Sg3~2100Hz)
- no moving articulators
  - ➝ fairly constant for a given speaker
- SGRs can distort spectral peaks of formants [1]
  - ➝ speakers avoid putting vowel formants in these regions, [2]
- formants are thought to be free to vary only within the frequency bands defined by the subglottal resonances
- recent studies for a few languages
  - [2], American English – relation between F2-Sg2
  - [4], English-Spanish bilingual – speaker normalization
  - [5], German, [6], Korean – relation between F1-Sg1, F2-Sg2
- ***hypotheses*: SGRs are natural divisions between +/- values of several distinctive features for Standard Hungarian:**
  - ***1) Sg1 is a boundary between low and non-low vowels***
  - ***2) Sg2 is a boundary between front and back vowels***
  - ***3) Sg3 is a boundary between front unrounded and front rounded vowels***

Background: the vowel system of Hungarian

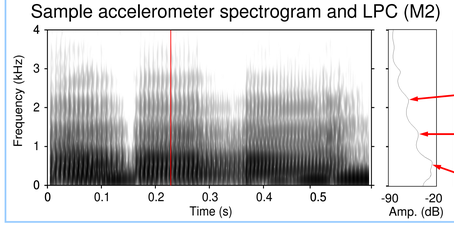


- 7 short vowels and 7 long vowels (all monophtongs)
- phonologically: paired together
- phonetically: quality differences between short / long versions low vowel pairs

Methods: recordings & measurements

- 2 male and 2 female adult native speakers of Standard Hungarian
- utterances: "ɔC_1VC_2ɔ" nonsense words
  - C_1: [b, d, g], C_2: [b] for males, [d] for females
  - V: target, all Hungarian vowels ([ɔ,aː,o,oː,u,uː,ɛ,eː,i,iː,ø,øː,y,yː])
  - 9-15 utterances per vowel
- simultaneous microphone and accelerometer recordings
- measurements:
  - first 3 formants of each vowel at midpoint, from microphone signal
    - automatic measurements using Praat + manual correction
  - first 3 subglottal resonances, from the accelerometer signal
    - manual measurements (25 times for each SGR/speaker)

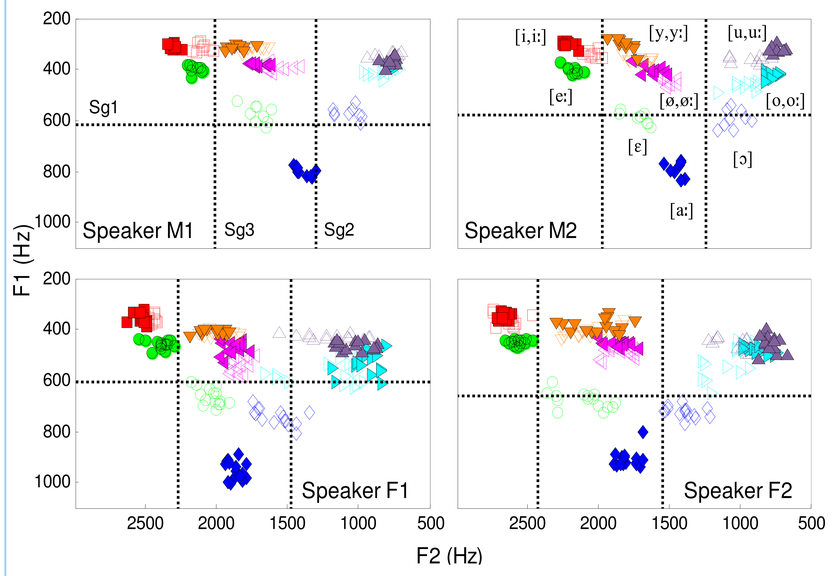
Sample accelerometer spectrogram and LPC (M2)



Medians of SGRs (Hz)

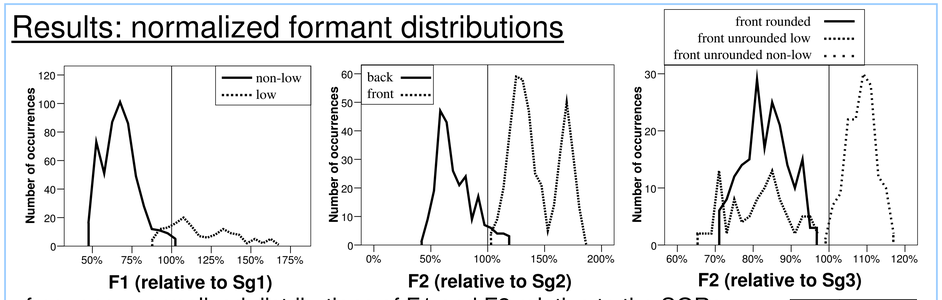
	M1	M2	F1	F2
Sg3	2015	1974	2268	2426
Sg2	1301	1235	1478	1543
Sg1	617	577	607	662

Results: vowel formants vs. subglottal resonances



- vowel spaces are clearly divided by the subglottal resonances
  - Sg1 (horizontal line) separates low and non-low vowels
  - Sg2 (right vertical line) separates front and back vowels
  - Sg3 (left vertical line) divides the front unrounded non-low vowels [i, iː, eː] from their rounded counterparts [ø, øː, y, yː]
- some exceptions from these rules
  - speaker M1: vowels [ɛ] and [ɔ] have F1 values lower than Sg1
  - speakers M2 and F2: F1 values of [ɛ] very close to Sg1
- potential explanations
  - resonances in the accelerometer signal are not as clean as formants in the microphone signal
  - speaker M1: strong coupling between the vocal tract and the subglottal system
  - co-articulation with consonants
  - SGRs do not separate low and non-low vowels as cleanly as hypothesized
- vowel [ɔ] : differences among speakers
  - speakers M1, M2 and F2: F2 lower than Sg2; speaker F1: F2 higher than Sg2
  - similar findings in other languages
    - German, [5], F2 of [a] categorically above / below Sg2, depending on the speaker
    - Korean, [6], F2 of [a] dependent on neighbouring consonant place of articulation

Results: normalized formant distributions



- frequency-normalized distributions of F1 and F2 relative to the SGRs, for all four speakers
- e.g. left figure: raw F1 values of each speaker were normalized with respect to Sg1 and then pooled together
- category separations are clearly visible as noted earlier

	correct
Sg1	94,94%
Sg2	97,47%
Sg3	99,77%

Results: ROC analysis

- a range of frequencies optimally separate the different categories of vowels
- out of 12 cases, median SGRs are
  - in 6 cases, within the optimal range
  - in 4 cases, within 1 S.D. of the optimal range
  - in 2 cases, out of the optimal range
    - ➝ discussed earlier (vowels [ɛ] and [ɔ])

Discussion, conclusions

- ***Sg2 is a reliable boundary between front and back vowels in Hungarian***
- ***Sg3 is a reliable boundary between front rounded* [ø, øː, y, yː] *and front unrounded non-low* [i, iː, eː] *vowels in Hungarian***
  - ***vowel* [ɛ] *is an exception, differs from* [eː] *in quality and quantity***
- ***Sg1 may be a boundary between low and non-low vowels in Hungarian***

- implications of the results
  - understanding phonological distinctive features
  - potential applications in speech technology: speaker normalization, speaker recognition

Key references

[1] Stevens, K. N., *Acoustic Phonetics*, Cambridge, MA: MIT Press,1998.
[2] Lulich, S. M., "Subglottal resonances and distinctive features,"J. Phon., doi:10.1016/j.wocn.2008.10.006, 2009.
[3] Gray, H., *Anatomy of the human body*. Philadelphia:Lea & Febiger, 1918.
[4] Wang, S., Alwan, A., Lulich, S. M., "Speaker normalization based on subglottal resonances," in *Proc. ICASSP*, 4277–4280, 2008.
[5] Madsack, A., Lulich, S. M., Wokurek, W., Dogil, G., "Subglottal resonances and vowel formant variability: A case study of high German monophthongs and Swabian diphthongs," in *Proc. LabPhon*, 11:91–92, 2008.
[6] Jung, Y., "Subglottal effects on the vowels across language: Preliminary study on Korean," JASA 125:2638, 2009.

Acknowledgements

- T. G. Cs. and T. B. were supported by the Hungarian government under contracts NKFP 2/034/2004, Jedlik OM-00102/2007 and TÁMOP-4.2.2-08/1/KMR-2008-0007
- Zs. B. was supported by HNRF PD050018 grant
- T. G. Cs. was partially supported by the Bizáki Puky Péter Memorial Foundation