



Introduction

Factors influencing the fundamental frequency (f0) across lifespan:

- physiological changes due to primary and secondary aging [1]; speech style, emotional background, sociolinguistic factors, etc.
- The most and heaviest aging-induced changes appear in elder ages, however, it starts in early adulthood, 30-35 years [2]

Cross-sectional studies: male speakers: f0 decreases to 50 years of age – after that: f0 rises (e.g., [3], summarized data: [4])

Longitudinal:

- relatively constant f0 values in 10 young adult subjects in 7 years periods following puberty [5] ↔ decrease in f0 values in young ages of the subjects, too [6], [7]
- significant deterioration of the acoustic voice signal & f0 increase in a 5 years period in smoking, this seemed reversible for males who stop smoking [8]

Most of the previous longitudinal studies examine:

- a low number of speakers; reading aloud or prepared speech, e.g., Christmas speech by Elisabeth II [6], [7]; sustained Vs and text fragments [8]

The aim of the study

Longitudinal analysis of spontaneous speech (SPSP), interpreted speech (ISP: summarizing a "listening-to-task") and reading aloud (RA) in decade

Hypotheses

- differences can be detected in the f0 in male subjects speech in the period of a decade
- differences can be detected in the f0 in male subjects speech between RA, SPSP & ISP

References

[1] Busse, E. W.: General Theories of Aging. In: Copeland, J. R. M., Abou-Saleh, M. T., Blazer, D. G. (ed.) Principles and Practice of Geriatric Psychiatry, Second Edition. Wiley-Blackwell (2002)

[2] Lalley, P. M.: The aging respiratory system-Pulmonary structure, function and neural control. Respiratory Physiology & Neurobiology 187, 199–210 (2013)

[3] Stathopoulos, E. T., Huber, J. E. & Sussman, J. E. Changes in Acoustic Characteristics of the Voice Across the Life Span: Measures From Individuals 4–93 Years of Age. J. Speech Lang. Hear. Res. 54, 1011–1021. (2011)

[4] Baken, R.J. The Aged Voice. A New Hypothesis. Journal of Voice 19(3), 317–325. (2005)

[5] Fouquet, M., Pisanski, K., Mathewson, N., Reby, D., 2016 Seven and up: individual differences in male voice fundamental frequency emerge before puberty and remain stable throughout adulthood. R. Soc. opensci. 3: 160395

[6] Harrington, J., Palethorpe, S. & Watson, C. I.: Age-related changes in fundamental frequency and formants: a longitudinal study of four speakers. In Proc. of Interspeech 2007, 2753–2756. (2007)

[7] Reubold, U., Harrington, J., Kieber, F.: Vocal aging effects on F0 and the first formant: A longitudinal analysis in adult speakers. Speech Communication 52, 638–651 (2010)

[8] Verdonck-de Leeuw, I. M., Mahteu, H. F.: Vocal aging and the impact on daily life: a longitudinal study. J. Voice 18(2), 193–202 (2004)

[9] Neuberger, T., Gyarmathy, D., Grácz, T. E., Horváth, V., Gósy, M., Beke, A.: Development of a large spontaneous speech database of agglutinative Hungarian language. In: Proc. of the 17th International Conference, TSD 2014, September 8–12, 2014, Brno. 424–431. (2014)

[10] Grácz, T. E., Huszár, A., Krepsz, V., Száraz, B., Damásdi, N., Markó, A.: Longitudinális korpusz magyar felnőtt adatközlőkről. [Longitudinal corpus of Hungarian adult speakers]. In: Berend, G., Gosztolya, G., Vincze, V. (szerk.) XVI. Magyar Számítógépes Nyelvészeti Konferencia, Szeged, 2020. január 23–24. 103–114.

[11] Boersma, P., Weenink, D.: Praat: doing phonetics by computer. http://www.fon.hum.uva.nl/praat/download_win.html

[12] R Core Team (2018). R: A Language and Environment of Statistical Computing. R Foundation for Computing, Vienna.

[13] Harrell, F. E. & Dupont, C. (2017) Hmisc: Harrell Miscellaneous. R package version 3.16.0.

[14] Bates, D., Mächler, M., Bolker, B. & Walker, S. (2015) Fitting Linear Mixed-Effects Models Using lme4. Journal of Statistical Software 67, 1–48.

[15] Kuznetsov, A., Brockhoff, P. B. & Christensen, R. H. B. (2017). lmerTest Package: Tests in Linear Mixed Effects Models. Journal of Statistical Software 82, 1–26.

[16] Akaike, H. (1974). A new look at the statistical model identification. IEEE Transactions on Automatic Control 19(6), pp. 716–723.

[17] Quené, H. (2013). Longitudinal trends in speech tempo: The case of Queen Beatrix. JASA 133, EL452. 452–457.

Methodology

1st rec: age between 19 and 40 years (mean: 27 ys, sd: 5 ys); BEA database [9]

2nd rec: age between 29 and 51 years (mean: 38 ys, sd: 5 ys); (Hungarian longitudinal corpus) [10]

Circumstances of the recording sessions: soundproof room, cardioid condenser microphone, 44,1 kHz sampling, 16 bit

Participants: native speakers of Hungarian without any speech or hearing deficits; 13 male subjects

Three speech styles: (i) a spontaneous speech (SPSP): interview about job, family, hobbies in a quasi-monological way, (ii) interpreted speech (ISP): summarizing an audio recording of a scientific outreach text, (iii) reading aloud (RA) of a(nother) scientific outreach text

Labeling: manually in Praat [11]: each interpausal unit was labeled (between two consecutive pauses or silences of any length and any type) that did not include noise/overlapping speech/backchannel/etc.

F0 analysis: 50 ms windows with 25 ms overlaps, f0 minimum and maximum: 50 to 300 Hz

The values outside the 1.5 * of the IQR range of each speech sample (/speaker/age/speech stlye) were eliminated.

Analysed features:

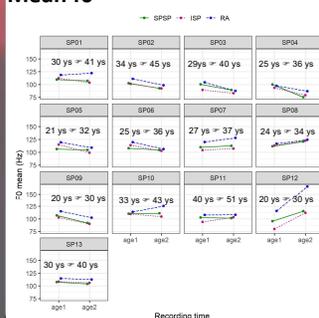
- F0 mean in Hz; f0 mean change calculated in semitones ($f0_{age2}$ to $f0_{age1}$)
- Global f0 range: the highest f0 value in semitones to the lowest (/speaker/age/speechstyle); global f0 range change: $value_{age2} - value_{age1}$
- Local f0 range: the highest f0 value in semitones to the lowest per in each interpausal unit (/speaker/age/speechstyle); global f0 range change: $value_{age2} - value_{age1}$

Statistics: R [12]; Hmisc [13], lme4 [14], lmerTest [15]

- LMM: random intercept for speakers, factors: speech style and age – added one by one & with and without random slope – picking the one with the lowest AIC-number [16]

Results

Mean f0



Change	Number of speakers		
Age2 to age1	SPSP	ISP	RA
Decr. min. 2 sts	4	8	6
Incr. min. 2 sts	2	3	4

	Mean & SD of change (semitones)
SPSP	-0.62±1.96
ISP	-0.43±2.35
RA	-0.17±2.44

The best fitting model:

Factors: age & speech style with random slopes for both, however, only the speech style had significant effect on the f0 ($F(2, 12.156) = 8.713, p = 0.004$).

Global f0 range

Age2 to age1	Number of speakers		
	SPSP	ISP	RA
+ min. 2 sts	1	7	5
- min. 2 sts	12	1	4

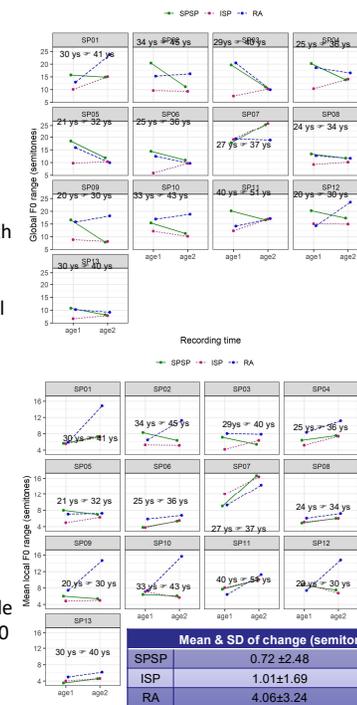
	Mean & SD of change (semitones)
SPSP	3.95±4.41
ISP	2.00±2.61
RA	-0.35±5.62

The best fitting model: factors: age & speech style with random slopes for both & interaction of these had a significant effect on the overall f0 range ($F(2, 47.998) = 8.890, p < 0.001$).

Local f0 range

Age2 to age1	Number of speakers		
	SPSP	ISP	RA
+ min. 2 sts	7	9	11
- min. 2 sts	4	2	0

The best fitting model: Factors: age & speech style with random slopes for both, however, only the speech style had significant effect on the f0 ($F(2, 12.372) = 4.6788, p = 0.031$).



Discussion & conclusions

F0 change varies among speech styles. F0 decrease is more typical than increase. Increase in SP12: no specific reason (= no disease or environmental effects)

Global f0 range decreased in SPSP, increased in ISP, & showed largely speaker-dependent pattern in RA.

Local f0 range increased in most speakers with some speech style dependency.

SPSP is the least formal both towards the addressee and in its topic. Also the speech planning processes are different across the speech styles thus this has larger effect than age. Also, further factors have to be considered in the interspeaker variability: [6] and [17] also suggested that not only the subject's age-related physiological changes but also further factors (e.g., sociolinguistic factors, like the addressee) have to be taken into consideration. Also a decade is a short time period for large changes in young and healthy speakers

Speaker identification may be affected by changes undergoing in a decade in the subject's speech characteristics.