

## **Fundamental frequency patterns: The factors of age and speech type**

*Alexandra Markó, Judit Bóna*

Department of Phonetics, Eötvös Loránd University, Budapest

marko.alexandra@btk.elte.hu, bona.judit@btk.elte.hu

### **Abstract**

This paper aims at investigating the effect of aging on speech in the case of two speech types. Two kinds of speech material were used in the experiments, read text and spontaneous narratives of 7 old and 7 young females. Fundamental frequency patterns and voice characteristics were analysed in both speech samples. Results show that (i) the F<sub>0</sub>-values had a wider range with older than with younger subjects in spontaneous speech, while (ii) the F<sub>0</sub>-values had a wider range with younger than with older subjects in reading aloud, and (iii) the frequency of glottalization is subject (and not age) dependent.

### **1. Introduction**

As a person progresses in age, the whole organism becomes older: a fact that affects the person's speech, too. Human voice production, and speech production in general, undergoes significant changes after 60 years of life; 'adult voice' turns into 'old voice' [1]. The aging process affects all speech organs. The pulmonary capacity decreases, resulting in a decrease of voice intensity (lung-power) and in a shortening of time for which speech sound production can be sustained [1]. The vocal cords and the laryngeal muscles lose some of their elasticity, making the fundamental frequency higher [2], [3]. The literature claims that the range of voice becomes narrower: as opposed to the 2 octaves of the average adult voice, elderly people's speech is characterized by slightly more than a single octave. Their voice becomes trembling. The movements of the tongue become more difficult, resulting in less accuracy of articulation. And the slackness of the soft palate may give rise to a nasalized quality in old people's speech [1].

Some of the most conspicuous changes are those of the clearness of voicing and the typical fundamental frequency values, possibly resulting in changes in intonation structures. Male speakers' F<sub>0</sub> rises while female speakers' F<sub>0</sub> does not change or becomes lower [4], [5], although some older speakers may have a 'younger' fundamental frequency and some younger speakers may have an 'older' one [6]. Individual differences are emphasized by a number of researchers. For instance, [7] analysed F<sub>0</sub> values of native English speakers over 100 years of age and showed that both average fundamental frequencies and F<sub>0</sub> minima are significantly lower with old speakers than with young ones; on the other hand, [8] attested higher F<sub>0</sub> values than usual in a case study involving a 105-year-old female subject. Other authors emphasize that F<sub>0</sub> changes in (Japanese) women are more extensive than in (Japanese) men [5].

The literature furthermore claims that F<sub>0</sub> variability is more extensive in old people's speech than with young speakers [7], [9], [10] and that old speakers produce more melisms in

reading aloud [11]. There is also a difference in terms of voice quality. Although several studies claim that there is no significant difference in jitter and shimmer values between the two age groups (e.g., [9]), most authors have attested increases in both acoustic parameters with growing age [11], [12], [13]. In native Hungarian female speakers, there are significant differences in both jitter and shimmer between young and old speakers, too; in particular, the voice of older speakers is 'more irregular' than that of younger ones [6]. The largest differences between the two groups were found in harmonics-to-noise ratio: this value is lower with older speakers, that is, their speech is 'noisier' than that of young speakers [6, 14].

The studies mentioned so far analyse acoustic parameters of voice production and intonation either in vowel articulations or in brief speech samples and disregard other factors such as glottalization, the direction of pitch courses, or the effect of diverse speech types. The latter is an important factor as several studies confirmed (for Hungarian, too) that the intonation structures of reading aloud and spontaneous speech differ [15], [16].

In the present study, the effect of aging on intonation is investigated in two different speech types: reading aloud and spontaneous speech production. Our hypothesis is that both factors (age and speech type) influence speech voice (in terms of voice quality and F<sub>0</sub> variation). The following differences are assumed to occur between reading and spontaneous speech:

- (i) the pitch range of spontaneous speech is wider since the planning of its suprasegmental structure is accidental and there may be more (meaningless) pitch variation in it;
- (ii) ends of breath groups in spontaneous speech often involve a rising contour, as in monologues this serves the indication that a continuation is to be expected [17].

The following differences between old and young speakers are predicted:

- (i) the F<sub>0</sub> range of older speakers is narrower, and their tonal interval values are lower, in accordance with data found in the earlier literature;
- (ii) older speakers more often exhibit non-modal phonation (e.g., glottalization, devoicing, or whisper). Glottalization (irregular voicing) may involve less air expelled from the lungs, hence lower subglottal pressure [18]; it can also be caused by less normal opening and closing of the vocal cords [19]. As older speakers are characterized by decreased lung capacity and weakening muscular tension [1], we have assumed that their speech will exhibit more glottalization, a phenomenon that is related to muscular activity. In addition, we hypothesise that the divergent operation of the vocal cords may cause involuntary devoicing and whisper in old people's speech;
- (iii) conventional beliefs sustain that breath group final pitch contours tend to be descending/falling in older people's speech but rising (or level) in that of young speakers [16].

## 2. Material, methods, subjects

For the present study, we selected 7 old (between 70 and 80, average: 73.9 years) and 7 young speakers (between 20 and 32, average: 26.6 years) from the Hungarian spontaneous speech database (BEA, cf. [20]). We analysed their speech production in two speech types: the reading of a text of popular science, 13 sentences long, and a spontaneous monologue in which the subjects talked about their hobbies and jobs. The reading technique of both age groups was adequate; none of the speakers had any difficulty in converting the printed text into speech. The durations of the samples can be seen in Table 1.

	Old	Young	Total
Read	17'31"	15'04"	32'35"
Spontaneous	20'46"	18'22"	39'08"
Total	38'17"	33'26"	71'43"

Table 1: Duration of the speech samples.

All samples were annotated by Praat 5.1 [21] at the breath group level; roughly 100 breath groups per subject were analysed. In all breath groups, we determined the maxima and minima of  $F_0$ , and then we calculated the intervals of each breath group and the overall ambit ( $F_0$  range) of the whole sample of each subject ( $F_{0max}/F_{0min}$ ). Furthermore, we analysed the group final realizations of melody in both speech types. We established the number of syllables that contained at least one glottalized, devoiced or whispered realization of a sonorant. Finally, we examined how age and speech type affected changes of fundamental frequency. The data were submitted to statistical analysis by SPSS 15.0 (one-way ANOVA and Tukey's post hoc test).

## 3. Results

$F_0$  range (or ambit) is a frequency domain defined by the maximal and minimal  $F_0$  values of a person's total speech production; it characterizes the speaker and the speech type, too (as former studies have established, cf. [22], [23]). We hypothesized that the  $F_0$  range of spontaneous speech was wider than that of reading aloud. As can be seen in Figure 1, the effect of age on  $F_0$  range was insignificant in the case of our subjects; the pitch domain used by a speaker is an individual characteristic. The two speech types nevertheless did have an effect, and age also played a role: it appeared to be determined by age how exactly  $F_0$  range was affected by speech type. In young subjects' speech, reading was invariably realized with a wider  $F_0$  range, while all old subjects (except one) used a wider range of fundamental frequency in their spontaneous speech.

The boxplot also shows that outliers (or extreme values, represented by numbered circles and asterisks in the figure) invariably occurred in spontaneous speech. While reading could be characterized by a compact set of realizations for all speakers, spontaneous speech could not. Except for three old informants (O1, O2, and O3), all speakers produced data between 200 and 400 Hz that lie outside the characteristic range of that speaker's spontaneous speech. The reasons for this must probably be seen in the role of chance in speech planning with respect to the suprasegmental structure of spontaneous speech, on one hand, and in the emotional surplus that may

occur in spontaneous speech, on the other (the text that had to be read did not give occasion for the expression of emotions).

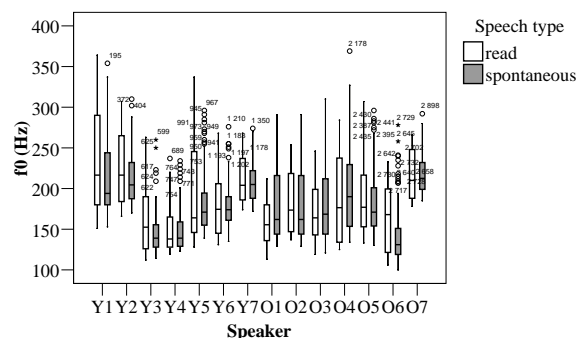


Figure 1:  $F_0$  ranges in the two speech types split by speaker (Y = young speakers, O = old speakers).

We also calculated pitch intervals (the highest  $F_0$  value divided by the lowest  $F_0$  value of the given breath group). Figure 2 shows these as a function of age and speech type. In both age groups, the average values of pitch intervals were higher for reading: for young speakers, the mean was 1.48 (s.d.=0.26); for old speakers, it was 1.45 (s.d.=0.22). The narrower intervals of spontaneous speech yielded 1.26 (s.d.=0.21) for young subjects and 1.36 (s.d.=0.29) for the old ones. The statistics gave a significant result across these four data sets ( $p < 0.001$ ;  $F(3, 1472) = 53.110$ ); the post hoc test showed that reading by old vs young subjects did not involve a significant difference of interval values, while all other comparisons revealed significant differences ( $p < 0.001$ ). This meant that reading aloud was characterized by more uniform intonation structures; such effect may partly be due to the nature of the written text itself, where segmentation is predetermined by punctuation and, consequently, suprasegmental structure is planned in advance. Actually all these factors did not apply to spontaneous speech. We should also mention the fact that the read speech has been elicited with respect to the same text for all the speakers, while spontaneous speech is – by definition – different across speakers. Then the presence of more prosodic variation in spontaneous speech could in principle be explained by the presence of more segmental, lexical and pragmatic variation, compared to read speech.

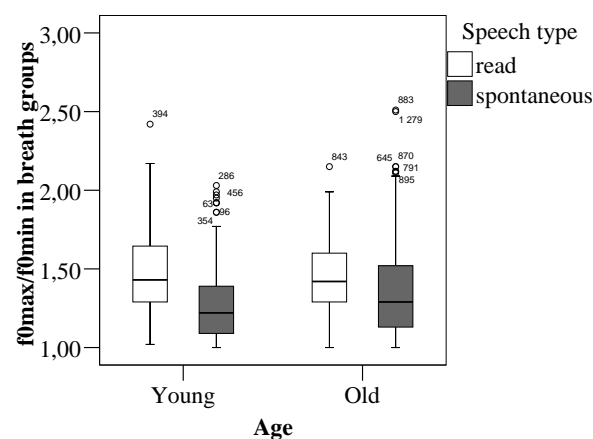


Figure 2: Pitch intervals of breath groups in the two speech types and in the two age groups.

We analysed the material with respect to voice quality. Deviations from modal voice were most often represented by glottalization, whereas the ratio of whispered syllables and those involving devoiced segments, taken together, did not reach 1% in the age groups' material. Glottalization was more frequent in spontaneous speech than in reading for both age groups, although the difference was not significant (Figure 3). In young subjects' reading, 9.96% of the syllables were partly or wholly glottalized; the same figure for old subjects is 13.80%. In young subjects' spontaneous speech, the ratio of glottalized syllables was 13.91%, while in the speech of old subjects, it was 16.45%. That is, we found relatively more glottalized syllables in the old speakers' material, compared to the young speakers' speech, for both speech types.

On the other hand, we found significant differences between the age groups. Young subjects glottalized 10.06% of the syllables on average in reading (s.d.=6.69%), while old subjects glottalized 13.94% of the syllables in the same task (but s.d. was 11.41% here). A similar tendency was observed in spontaneous speech, with young speakers producing 14.11% glottalization (s.d.=5.55%), and old speakers producing 17.27% glottalization (s.d.=12.94%). Old subjects exhibited significantly more inter-speaker variation (Figure 4). One of them (80 years old, the oldest in the group) glottalized more than 40% of her syllables in reading and over 50% of her syllables in spontaneous speech.

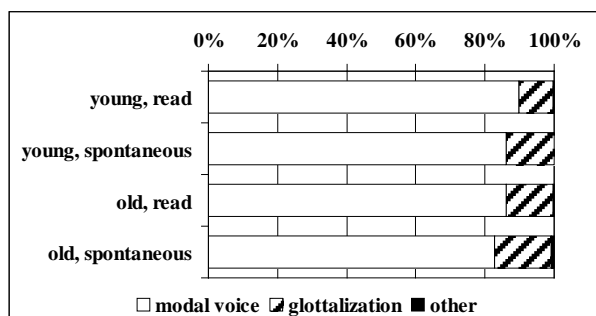


Figure 3: Voice quality data as a percentage of the total number of syllables.

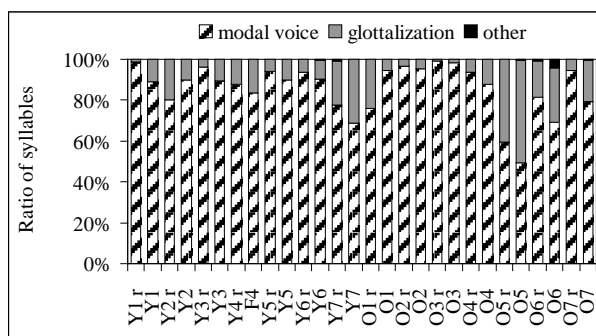


Figure 4: Voice quality data by speaker (Y = young, O = old, r = reading).

It is a common belief that young speakers' speech (and especially that of young women) is characterized by final pitch rises and/or final steady pitches [16]. In order to check the validity of this claim, we analysed the melody (pitch contour) realizations of ends of breath groups by age and by speech type. Our results support the claim that breath group final

steady pitch is characteristic of young persons' spontaneous speech: 48.12% of their final contours in this speech type were steady. On the other hand, final rise was not more frequent in young people's speech than in old people's. In young subjects' reading, old subjects' reading, and old subjects' spontaneous speech, the proportion of steady, rising and falling pitch was very similar indeed (Figure 5). In these categories, final glottalization amounted to 21.74-22.43% of the occurrences, falling pitch was between 17.30-21.38%, steady pitch between 27.54-33.12%, rising pitch between 25.32-28.62%, and all other possibilities (including whispered, devoiced, and laughing-coloured ends of breath groups) occurred in 0.00-2.32% of all cases.

The relatively high ratio of glottalized endings in all groups of data is worth mentioning. On the basis of our results, it can be assumed that in spoken Hungarian – just like in other languages, e.g. American English [24], [25] – glottalization has a boundary marking function. Other investigations [16], [26], [27] seem to independently support this conclusion for Hungarian.

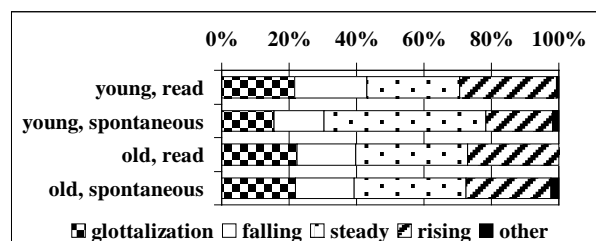


Figure 5: The intonation at the end of breath groups.

## 4. Conclusions

In this paper, we compared the properties of young and old female speakers' intonation in a corpus involving two speech types, reading aloud and spontaneous speech production. The most important results can be summarized as follows.

With respect to pitch range, our data did not confirm the claim in the literature that pitch range gets smaller with age: in both speech types, both young and old speakers' range was roughly an octave (that is, the ratio of minimal and maximal  $F_0$  values was 1:2). Standard deviation was also very similar across groups: the  $F_{0max}/F_{0min}$  ratio for speakers with the narrowest ambits fell between 1.49 and 1.65 in both speech types and both age groups, and the widest range was somewhere between 2.31 and 2.78.

The factor of age affected the differences between speech types in opposite ways: in the case of young speakers, reading was realized with a somewhat wider  $F_0$  range (2.13 vs 2.01), whereas in the case of old speakers,  $F_0$  range was wider for spontaneous speech than for reading (2.01 vs 2.35). Our first hypothesis (according to which the pitch range of spontaneous speech would be wider) was therefore only partially confirmed. One reason for the attested difference between age groups may be that younger speakers were more careful about expressive reading; they realized more variegated suprasegmental structures. Another reason may be that the spontaneous speech of older subjects involved more emotions (probably due to the memories of long-gone experiences as most of them were talking about their former jobs); this resulted in wider variability of  $F_0$  values and higher pitch interval values in the spontaneous speech of old subjects, compared to young people.

We also hypothesized that the ends of breath groups would be characterized by pitch rises in spontaneous speech more often than in reading; this, however, was not confirmed. Similarly, no confirmation came for the claim that young people's (spontaneous) speech would involve more pitch rises at the end of breath groups than old people's speech. On the other hand, steady-pitched ends of breath groups dominated in the spontaneous speech of young subjects.

In the production of older speakers there were more cases of phonation differing from modal voice (primarily, glottalization) than in the case of young speakers; however, the differences between the two groups were small. In terms of ratio of glottalization, individual differences within the old group of subjects were more marked than those between the two age groups. This means that, with growing age, voice production becomes more uncertain for some speakers than for others, that is, the physiological effects of aging may differ across speakers.

The data reveal that speech type affects the parameters studied here more than age does. This suggests that the comparison of reading and spontaneous speech must be taken into account in forthcoming studies. The results of the present paper, in accordance with a number of other studies, draw our attention to the fact that age in itself does not account for all peculiarities of the fundamental frequency of speech; these largely are determined on an individual basis at any age. Longitudinal studies might provide answers to the question of how growing age modifies the characteristics of intonation, and whether other voice factors are influenced by such changes.

## 5. Acknowledgements

The writing of this paper was supported by the Bolyai János Research Grant. We would like to thank Péter Siptár for his help.

## 6. References

- [1] Balázs B. 1993. Az időskori hangképzés jellemzői [Properties of voice production in older speakers]. *Beszédkutatás* '93. 156-165.
- [2] Watson P.J. and B. Munson. 2007. A comparison of vowel acoustics between older and younger adults. In Trouvain J. and W.J. Barry (eds.) *Proceedings of the 16th International Congress of Phonetic Sciences*. Saarbrücken, Germany. 561-564.
- [3] Torre P. and J.A. Barrow 2009. Age-related changes in acoustic characteristics of adult speech. *Journal of Communication Disorders* 42. 324-333.
- [4] Russel A., L. Penny and C. Pemberton. 1995. Speaking fundamental frequency changes over time in women: A longitudinal study. *Journal of Speech and Hearing Research* 38. 101-109.
- [5] Nishio M. and S. Niimi 2008. Changes in Speaking Fundamental Frequency Characteristics with Aging. *Folia Phoniatrica et Logopaedica* 60. 120-127.
- [6] Bóna J. 2009. Az idős életkor tükröződése a magánhangzók ejtésében [The signs of old age in vowel production]. *Beszédkutatás* 2009. 76-87.
- [7] Awan S.N. and P.B. Mueller. 1992. Speaking fundamental frequency characteristics of centenarian females. *Clinical Linguistics and Phonetics* 6. 249-254.
- [8] Max L. and P.B. Mueller. 1996. Speaking F0 and cepstral periodicity analysis of conversational speech in a 105-year-old woman: Variability of aging effects. *Journal of Voice* 10. 245-251.
- [9] Brown W.S., R.J. Morris and J.F. Michel. 1989. Vocal jitter in young adult and aged female voices. *Journal of Voice* 3. 113-119.
- [10] Gorham-Rowan M.M. and J. Laures-Gore. 2006. Acoustic-perceptual correlates of voice quality in elderly men and women. *Journal of Communication Disorders* 39. 171-184.
- [11] Benjamin B.J. 1981. Frequency variability in the aged voice. *The Journal of Gerontology* 36. 722-726.
- [12] Biever D.M. and D.M. Bless. 1989. Vibratory characteristics of the vocal folds in young adult and geriatric women. *Journal of Voice* 3. 120-131.
- [13] Orlikoff R.F. 1990. The relationship of age and cardiovascular health to certain acoustic characteristics of male voices. *Journal of Speech and Hearing Research* 33. 450-457.
- [14] Ferrand C.T. 2002. Harmonics-to-noise ratio: An index of vocal aging. *Journal of Voice* 16. 480-487.
- [15] Beke A. 2008. A felolvasás és a spontán beszéd alaphangszerkezetének vizsgálata [A study of the fundamental frequency structure of reading and spontaneous speech]. *Beszédkutatás* 2008. 93-107.
- [16] Markó A. 2009. Stigmatizált hanglejtésforma a spontán beszédben [A stigmatized intonation contour in spontaneous speech]. *Beszédkutatás* 2009. 88-106.
- [17] Markó A. 2007-2008. Szakaszvégi dallamrealizációk a spontán beszédben [Breath group final melody realisations in spontaneous speech]. *Nyelvtudomány III-IV*. 2007-2008. 129-140.
- [18] Blomgren M., Y. Chen, M.L. Ng and H.R. Gilbert. 1998. Acoustic, aerodynamic, physiologic, and perceptual properties of modal and vocal fry registers. *Journal of the Acoustical Society of America* 103. 2649-2658.
- [19] Slifka J. 2000. *Respiratory constraints on speech production at prosodic boundaries*. PhD dissertation. Cambridge, Massachusetts: MIT.
- [20] Gósy M. 2008. Magyar spontánbeszéd-adatbázis – BEA [BEA: A Hungarian spontaneous speech database]. *Beszédkutatás* 2008. 194-207.
- [21] Boersma P. and D. Weenink. 2009. *Praat: doing phonetics by computer (Version 5.1)*. [http://www.fon.hum.uva.nl/praat/download\\_win.html](http://www.fon.hum.uva.nl/praat/download_win.html).
- [22] Fónagy I. and K. Magdics. 1967. *A magyar beszéd dallama* [The melody of Hungarian speech]. Budapest: Akadémiai Kiadó.
- [23] Wacha I. 1974. Az elhangzó beszéd főbb akusztikus stílus kategóriáiról [On some major stylistic categories of speech]. *Általános Nyelvészeti Tanulmányok* X. 203-216.
- [24] Henton C.G. and Bladon A. 1987. Creak as a sociophonetic marker. In Hyman L.M. and C.N. Li (eds.). *Language, speech and mind: Studies in honour of Victoria A. Fromkin*. London: Routledge. 3-29.
- [25] Redi L. and S. Shattuck-Hufnagel. 2001. Variation in the realization of glottalization in normal speakers. *Journal of Phonetics* 29. 407-429.
- [26] Markó A. 2010. A prozódia szerepe a spontán beszéd tagolásában [The role of prosody in the segmentation of spontaneous speech]. *Beszédkutatás* 2010. 82-99.
- [27] Markó A. 2011. A glottalizáció határjelző szerepe a felolvasásban [The boundary marking function of glottalization in reading aloud]. *Beszédkutatás* 2011. 31-45.