MASSIMO PETTORINO
ANTONELLA GIANNINI
Università degli Studi di Napoli L'Orientale, Napoli, Italy
mpettorino@unior.it
agiannini@unior.it

The voice and aging

Introduction

The relationship between the speaker's age and his voice has been studied in many experimental researches. It has already been proved that, both on the articulatory [Linville 2004; Awan 2006] and acoustic [Hollien and Ship 1972; Pegoraro-Krook 1988; Stoicheff 1981; Russell et alii 1995] and perceptive level [Horri and Ryan 1981; Hartman and Danhauer 1976; Schötz 2005], the voice changes in function of the age. The main reasons of these changes are due to anatomical and physiological variations involving not only lungs but also larynx and vocal tract. Growing old a lowering of the breathing functions occurs due to different reasons. Among them a muscle relaxation, a reduced vital capacity, a hardening of the vocal folds that become thinner, stiffer and less elastic, a laryngeal cartilage calcification and, finally, a hardening of the phonatory apparatus. The aging of the voice depends on many factors and it varies from one person to another. One of the weak points of this kind of studies is that the the speaker and the text are never the same. There are in fact two types of corpora: there is either the same text read by speakers of a different age, or a spontaneous speech uttered by differently aged speakers. However this method cannot guarantee that the results of the analyses are due exclusively to the speaker's age. In fact our speech varies according to many factors as, for instance, the contextual situation, the relationship with the interlocutor, formal/informal style, the subject of the conversation, read or spontaneous speech. Following a methodology already used in our previous works [Giannini and Pettorino 2007], we made an experiment: we asked a well known TV reporter, Piero Angela, to read the same text he read 40 years ago on the TV news. So, for the first time we could dispose of two identical corpora (same speaker, same situation, same text, same type of speech) but only one parameter, that is the age of the speaker.
The spectroacoustic analyses conducted on these two corpora pointed out the following articulatory and acoustic correlates of the aging:

- a lowering of the speech rate and of the articulation rate (about one syllable/second less);
- a lower fluency (three syllables less between one silence and the other);
- a slightly higher percentage of silence;
- a longer syllable;
- longer vowels in the stressed open syllables;
- the speech is more isochronic;
- an increase of silent pauses, both in number and in duration;
- shorter speech chains;
- a higher f_0;
- a larger tonal range;
- a wider vocalic area.

So, the close relationship between voice and speaker's age was confirmed. This explains why, on the perceptive level, a voice is sufficient to give to the listener an idea of the age of the speaker. Many studies investigated on the possibility of inferring the speaker's age only on the basis of a voice. The accuracy of the judgment depends on many factors, as for instance the type of speech (read/spontaneous), the speech rate, the listener's age, the speaker's dialect and so on [Hollien and Shipp 1972, Ptacek and Sander 1966, Linville 1987].

The purpose of the present study is to verify, on perceptive basis, with what degree of approximation is possible to identify the speaker's age and if there are evaluation differences due to the sex and the age of the listener. Besides, we intend to verify, through a spectrographic analysis, to what extent the Articulation Rate (AR) and the Speech Rate (SR) affect the identification of the speaker's age[1].

Materials and methods

The corpus used for this study has been obtained as follows. We chose, from a newspaper article, a sentence lasting about 7 seconds. The text, without punctuation, was read by 23 male italophone speakers, all coming from Campania region, divided into six different age ranges. The voices were recorded in a quiet but not soundproof room. We identified each speaker with a code, such as **134CS**, where **CS** stands for the initials of the speaker's name and surname, **1** stands for the first age range, **3** for the third height range and **4** for the fourth weight range (TAB. I). For this study we just consider the age data; we will deal with the height and weight data in a wider work concerning the speaker's identification.

[1] The **AR** is the ratio between the number of syllables and the articulated time. The **SR** is the ratio between the number of syllables and the total time.

age ranges	age	height	weight
1	16-25	1,50-1,59	40-49
2	26-35	1,60-1,69	50-59
3	36-45	1,70-1,79	60-69
4	46-55	1,80-1,89	70-79
5	56-65	1,90-1,99	80-89
6	66-75		90-99

TAB. I. *Age, height and weight ranges of the speakers*

The research consisted of two phases. In the first phase, through the use of *Goldwave* software, we created a single track composed by 10 pairs of voices. One of the pairs was included twice as a control group in order to evaluate the persuasion degree of the listeners' answers. The components of each pair differed from each other for one or more age ranges. Between the voices of each pair there was a 0.5 s pause.

The shift from a pair to the other was marked by a female guide voice. In the second phase, we have randomized in a new track 6 voices, taken from those used for the first experiment, together with 14 new voices. The corpus, thus obtained, has been the object of two different perceptive experiments made by 177 listeners, 90 female and 87 male, selected according to the age range as follows:
15 men and 16 women for the age range 1;
15 men and 15 women for the age ranges 2-4-5;
14 men and 15 women for the age range 3;
13 men and 14 women for the age range 6.
The track was listened in a quiet room and each listener had to his/her disposal headphones connected to a computer. The two tests were made in the same session. Each listener was given a form for each test. According to the sex, age, height and weight, we have assigned a code to each listener, conforming to the labelling method used for the speakers (see TAB. I) and differing only for an **F** (for female) or an **M** (for male) added before the code.

Experiment 1

In the first test for each voice pair the listener had to indicate the youngest voice. According to the given answers, each listener was assigned a score: 1 for a correct answer and 0 for wrong answer. The results of the first test are reported in TAB. II.

The average percentage of the correct answers corresponds, therefore, to 88%. More specifically, 76 subjects (equal to 43% of the listeners) did not make any mistakes, correctly recognizing the 10 pairs of the test, 45 listeners (the 25%) made just one mistake, 26 (the 15%) 2 mistakes, 11 (the 6%) and 12 (the 12%) made respectively 3 and 4 mistakes, 3 (the 2%) 5 mistakes and only one listener made 8 mistakes. Moreover, we have to remark that the control group (in bold in TAB. II) produced a higher number of correct answers at the second listening.
Pairs 5 and 6 are those that obtained the lowest score. A possible explanation could be that the two voices, differing from each other for just one or two age ranges, seemed perceptively more similar. Nevertheless, data are ambiguous at this regard because the answers relating to the ninth pair, where the two elements differed from each other for just one range, obtained a better result than those belonging to more distant ranges. Furthermore, different pairs with the same range difference (e.g. pair 4 and pair 7) obtained a different identification percentage. The perception of the age, thus, has to be connected to acoustically determined parameters such as the articulation rate (AR) and the speech rate (SR). In figure 1, for each pair, are reported the AR and the SR values.

pairs	speaker A	speaker B	difference of age ranges	correct answers	wrong answers	% of correct answers
1	244ML	534VL	3	163	14	92
2	**426MM**	**144CC**	**3**	**153**	**24**	**86**
3	335NC	623ND	3	158	19	89
4	534PS	146CV	4	154	23	87
5	433CS	326AG	1	140	37	79
6	335VB	123AM	2	129	48	73
7	235LL	634AG	4	168	9	95
8	**426MM**	**144CC**	**3**	**163**	**14**	**92**
9	244ML	334OA	1	160	17	90
10	623ND	144LI	5	163	14	92
	total			**1551**	**219**	**88**

TAB. II. *Results of the first test (in bold the control group)*

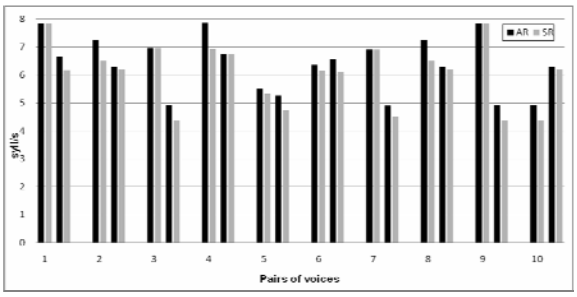

Fig. 1. *Articulation rate and speech rate of different pairs of voices*

As we can see pairs 5 and 6, pairs which obtained the lowest identification score, have a quite similar AR and SR. Nevertheless, we have to remark that pair 9, characterized by a difference of about 3 syllables/s (either for AR or SR) did not obtain the highest identification score. On this subject, data are not unambiguous and other parameters should be considered, such as pitch and fluency.

At this point we wondered if there was a relationship between the answers and listeners' sex and age. As regards the distinction between men and women, the outcome has shown that there are no differences in the perception of voice: men gave the 87% and women the 88% of correct answers. On the contrary, as regards the listeners' age it seems that there are important differences. As we can see in figure 2, listeners belonging to the first and the second age ranges are those who gave, in percentage, the highest number of correct answers (respectively 94% and the 91%), while those belonging to the sixth age range gave the lowest number of correct answers (77%).

These data are extremely useful because they allow us to identify and select listeners who are more suitable to the identification of the speaker's age, so the “ideal” listener should have an age included between 16 and 35 years.

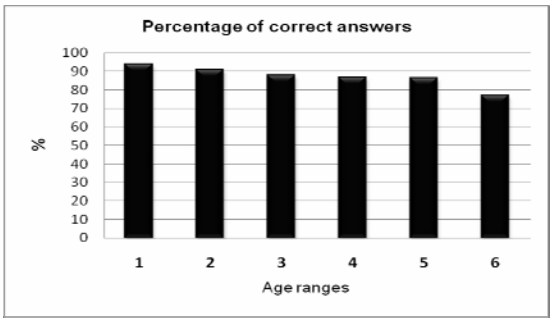

Fig. 2. *Percentage of correct answers for different age ranges of listeners*

Experiment 2

In the second test the listeners had to indicate, on a form, the age of each speaker. In figure 3 are reported the average values of the perceived and chronological ages. On the *y-axis* are reported the ages and on the *x-axis* are represented the speakers grouped according to the age range.

As we can see, the chronological age included in the first range is overestimated, while the voices of ranges 4, 5 and 6 are underestimated. The best results are those of the second and the third ranges. On the whole, the difference between chronological and perceived age is limited to just one range. The two repeated voices (335NC and 446 ADP) obtained quite the same result and this fact testify the non-randomness of the answers.

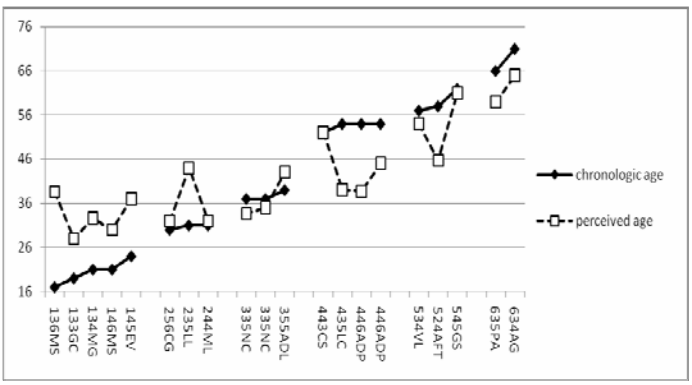

Fig. 3*. Perceived age versus chronologic age (average values)*

Conclusions

The results of the perceptive test allow us to draw the following conclusions:

1. the voice contains some acoustic traits related to the speaker's age and these traits can be perceptively recognized by the listener;
2. the comparative test, where the listener had to decide which one of the two voices was the youngest/oldest, shows that the identification is correct with a very high probability (88%);
3. younger listeners (16-35 years) recognize the age of the speaker better than the older ones; on the contrary, we have not noticed any differences as regards the listeners' sex;
4. the second test shows that the age of the speaker tends to be overestimated for the younger ranges (16-25) and underestimated for older ones (46 upwards).

On the basis of the results of this research, it seems that it is possible to recognize, with a good approximation, the age of the speaker on an exclusively perceptive basis. To arrange a procedure able to give the best results we should carry out further researches, proceeding as follows:

1. to choose listeners with an age included between 16 and 35 years;
2. to make a series of comparative tests where the same voice, whose age is unknown, is compared to voices belonging to different age ranges;
3. the results of the first series of tests will allow us to concentrate the analysis on a reduced number of age ranges. The second test will involve only the ranges thus identified and this will allow to reduce further on the age range of the speaker.

Such a procedure could be useful for several applications, as for instance the identification of the speaker's age in a forensic field, the protection of the children in web navigation, or in all those cases where the subject is not able to give his personal data (medical, social, immigrants welfare fields).

References

Awan, S. N, (2006), The aging female voice: acoustic and respiratory data, *Clinical Linguistics & Phonetics*, vol. 20, 2 & 3, April 2006, 171-180

Giannini A., Pettorino M. (2007), L'età della voce. In *La Fonetica Sperimentale: metodo e applicazioni, Atti del IV convegno Nazionale AISV*, Romito L., Galatà V. e Lio R. (a cura di), Università della Calabria, 3-5 dicembre 2007, EDK Editore SRL: RN, 2009, Vol. 4, pp. 165-178

Hartman, D. E. & Danhauer, J.L. (1976), Perceptual features of speech for males in four perceived age decades", *Journal of the Acoustical Society of America*, 59 (3), 713-715

Hollien, H. &Shipp, T. (1972), Speaking fundamental frequency and chronological age in males, *Journal of Speech and Hearing Research*, 15:155-159

Horri, Y. & Ryan, W. J. (1981), Fundamental frequency characteristics and perceived age of adult male speakers, *Folia Phoniatrica*, 33: 227–233.

Linville, S. E. (1987), Acoustic-perceptual studies of aging voice in women, *Journal of voice,* 1 (1), pp. 44-48

Linville, S.E. (2004), The Aging Voice, *American Speech-Language Hearing Association*, pp. 12-21

Pegoraro-Krook, M. I. (1988), Speaking fundamental frequency characteristics of normal Swedish Subjects Obtained by Glottal Frequency Analysis, *Folia Phoniatrica*, 40: 82-90

Ptacek, P. H., Sander, E. K. (1966), Age recognition from voice, *Journal of Speech and Hearing Research*, 9:273-277

Russell, A., Penny, L., Pemberton, C. (1995), Speaking Fundamental Frequency Changes Over Time in Women. A Longitudinal Study, *Journal of Speech and Hearing Research,* 38: 101-109

Shipp, T. & Hollien, H. (1969), Perception of the aging male voice", *Journal of Speech and Hearing Research*, 12:703-710

Stoicheff, M. L., (1981), Speaking fundamental frequency characteristics of nonsmoking female adults, *Journal of Speech of Hearing Research,* 24:437-441