Age and Rhythmic Variations. A Study on Italian.

Pettorino Massimo, Pellegrino Elisa

University of Naples "L'Orientale"

mpettorino@unior.it, epellegrino@unior.it

Abstract

Age-related variations in the speech production mechanism affecting acoustic cues such as pitch, speech rate, formant frequencies have been extensively investigated. Changes in speech rhythm and in utterance composition in terms of vowel and consonant portions are, instead, scarcely examined. Given the relevance of the vowel percentage (%V) in the utterance and of the interval between two vowel onset points (VtoV) for the perception of language rhythm, this study aims to investigate the relationship between rhythmic variations and speaker age. It also attempts to determine whether %V can be affected by different speech rates. Four young adult and four old speakers of Italian read four sentences of different length, at four different speech rates. The whole corpus was segmented in vocalic and consonantal portions and in VtoV intervals. The analysis results have shown that aged voice accompanies with significant increase in %V, despite speech rate variations. With advancing age, Italian speech tends to shift from the isosyllabic towards the isomoraic rhythmic pattern.

Index Terms: voice and aging, speech rhythm, speech rate

1. Introduction

The process of aging is generally associated with a number of changes in different domains, including the structure and mechanism of vocal tract [1]. The functional changes occurring in the respiratory, phonatory, and supralaryngeal systems from early childhood to old age alter the way human beings speak insomuch as listeners are able to judge speaker age fairly accurately from speech sample alone [2-6].

Over the years, age-related variations in the speech production mechanism have been extensively investigated. Stiffening of thorax, decreased lung capacity, and weakening of respiratory muscles are the most significant changes in the respiratory system [7]. With advancing age, the larynx also undergoes anatomic changes, including ossification and calcification of laryngeal cartilages and atrophy of vocal folds. Speech is affected also by changes in the supralaryngeal system [8]. Some researchers report of diminished accuracy of the lower lip and jaw when performing rapid movements, decreased lip strength and atrophy of facial, mastication and pharyngeal muscles [9], [10].

All these physical changes cause complex variations in the speech signal. However, recognizing the specific role played by aging in the voice changes is not always a clear cut. The findings in acoustic studies on speaker age, indeed, may be affected by different kinds of factors [11]:

- speaker-related factors (such as gender, race, weight, health and language, dialect, emotional state, attitude);
- speech-material-related factors (number and age distribution of the speaker, duration and speech type of the speech samples);
- methodological factors (recording, recording equipment and technique).

Despite the numerous variables that can affect the correlates of adult speaker voice, there is a general agreement among researchers on several acoustic features which vary with age (see [12] for a comprehensive overview on aged voice).

Several studies on pitch level have shown differences between male and female aging. After middle age, F0 rises substantially in men (about 35 Hz), while decreases in women (about 10-15 Hz). Amplitude stability also declines with aging, at least in men. Regardless of gender, instead, tremor and hoarseness increase because of a decline in F0 stability. [13], [14] and [15]. Other voice qualities generally linked with the aged voice are:

- increased jitter and shimmer [13], [16], [17];
- increased breathiness [13];
- lowering of formant frequencies [12], [13];
- altered vowel formant frequency patterns [13];
- altered resonance pattern [13];
- lengthening of vowels and stop consonants [3];
- lowering of speech rate [18].

Variations in speech rhythm and in utterance composition in terms of vocalic and consonantal portions are generally not taken into account in the research on the acoustic correlates of adult speaker voice.

2. Preliminary research on age-related rhythmic variations

Preliminary evidence of age-related rhythmic variations in Italian have emerged in a case study investigating the extent to which prosody and intonation affect listeners' ability to estimate the speaker age [19]. In that study the speaker-related variables, the speech material-related factors and the methodological variables mentioned above were controlled. A particular corpus of read speech was thus collected. In 2007, a 79-year old Italian anchorman, Piero Angela (henceforth PA), was asked to read a script that he had already read before in a 1968 TV news, acting as if he were hosting a real TV news broadcast. The recording was taken at RAI TV studios in Rome, in order to maintain the same communicative situation. Spectro-acoustic analyses were conducted on both corpora, the one uttered in 1967 at the age of 40, the other in 2007 at the age of 79.

The results have shown that the old speech, besides the variations that consistently appear in the literature to change with chronological age (wider tonal range and register, longer and more frequent silences, decreased articulation and speech rates), presented higher percentage of vowel portion (51%) than the young speech (46%) [19]. This specific variation deserves further investigation since the vowel percentage in the utterance (%V) is proved to be a reliable parameter enabling listeners to identify the rhythm of an utterance [20]. Studies carried out on a number of languages [21], [22] have

indeed shown that the three rhythmic groups (syllable, stress and mora-timed languages) differ considerably in vowel percentage and in another parameter, that according to [21] is ΔC (standard deviation of consonantal portions) and according to [23] is the interval between two consecutive vowel onset points (VtoV).

Stress-timed languages with their wider variety of syllable structure type, more complex consonantal clusters, higher frequency of CVC syllables and drastic reduction of unstressed vowels are characterized by lower %V (about 40-42%) and larger ΔC and VtoV values than syllable- and mora-timed languages. Mora-timed languages, characterized by a prevalence of CV syllables [24] accompany with the highest percentage of vocalic portion (about 53%) and the lowest degree of variability in terms of consonantal duration. Between these two extremes there are the syllable-timed languages with intermediate %V (about 44-46%) and ΔC/VtoV values (fig. 1).

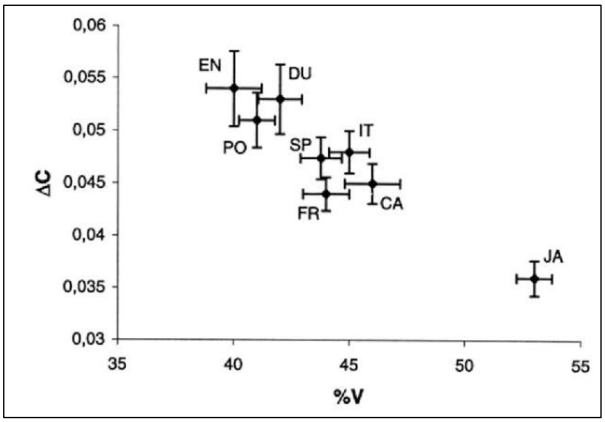

Figure 1: *Distribution of languages along the %V/ ΔC axis (from Ramus et al.* [21]*).*

In order to process the data concerning age-related variations of vocalic portion in the theoretical framework of rhythmic organization of languages, average %V and VtoV values obtained in Piero Angela's corpus (henceforth PA corpus) were compared to those resulting from the analysis of a multilingual corpus of TV news reading, recited speech, drugs advertising produced by speakers in their Forties [23] (fig. 2).

The figure 2 shows that %V and VtoV values of 40-year old PA overlap to those scored by Italian TV news broadcast ($p>0.05$). By contrast, the mean VtoV and %V values of the utterances produced by 79-year old PA are not statistically different from those scored by the Japanese anchorman ($p>0.05$). Thus, with advancing age, Italian speech increases in vocalic portion, shifting towards the isomoraic rhythmic pattern. The same age-related increase in %V was found in the speech produced by a 80-year old Japanese speaker. If this speech is compared to that of the 40 year old Japanese anchorman, %V increases from 50.7 to 55.4%.

The higher %V values scored by the old Italian and Japanese speakers are not divergent from results of studies on the rhythmic features of languages, since the speakers involved in the studies [21], [22], [23] were aged between 25 and 40.

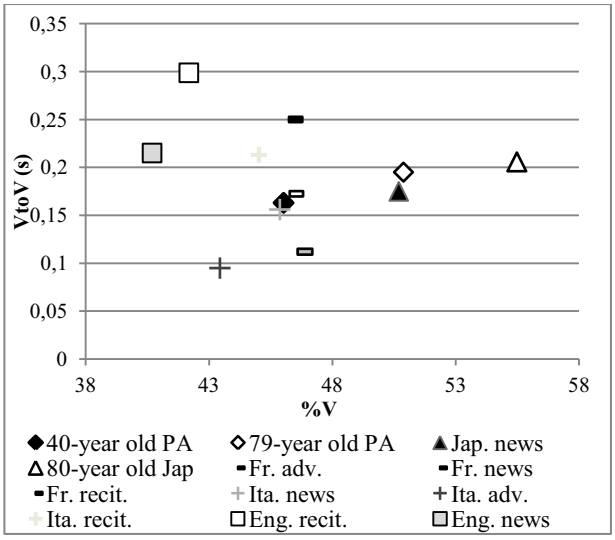

Figure 2: *%V and VtoV of different languages*.

3. The study

In the present study the correlation between speaker age and %V variations is examined in depth. The objective of the research is twofold: firstly, it is intended to figure out whether the rhythmic variations found in PA corpus were speaker- or age-dependent. Secondly, it attempts to determine whether, chronological age being equal, the vocalic portion of the utterance is affected by different speech rates.

To the purpose, eight Italian native speakers, equally distributed between the genders (4 M - 4 F) and two age groups (4 young adult – 4 old) were involved in the research. The four young adults ranged in age from 20 to 25, the four old were aged between 75 and 80.

Participants were instructed to utter four sentences of different length, from 20 to 35 syllables, displayed on a computer screen. The sentences were:

1. Il presidente ha ricevuto il ministro e ha accettato le dimissioni.
2. Il presidente ha ricevuto il ministro degli esteri e ha accettato le sue dimissioni.
3. Il presidente della repubblica ha ricevuto il ministro degli esteri e ha accettato le sue dimissioni.
4. Ieri mattina il presidente della repubblica ha ricevuto il ministro degli esteri e ha accettato le sue dimissioni.

To control the speech rate, participants were suggested to synchronize the performance of each sentence with the beginning and the end of the time interval for an arrow, displayed on the same computer screen, reaching the center of a target. This time interval lasted five seconds. This allowed to collect a corpus of sentences performed at four different articulation rates: about 4, 5, 6 and 7 syll./s. These rate values correspond in Italian to "slow", "medium", "fast" and "very fast" speech [25], [26], [27], [28].

The recordings were taken in single sessions in the silent room of the University of Naples L'Orientale, at 44.100 Hz sampling rate. The resulting corpus was therefore composed of 32 utterances (4 sentences * 8 speakers), each lasting about five seconds.

On the whole corpus spectro-acoustic analysis was carried out by means of Praat [29]. The signal was annotated in two

different tiers. In the first tier, the signal was segmented in VtoV intervals, in the second into vocalic and consonantal portions. The glides were segmented according to the rules adopted by Ramus et al. [21]: [w] and [j] were treated as consonants and the boundary was placed between the approximant and the vowel. Falling diphthongs were segmented in one or two vowels, depending on the spectro-acoustic characteristics of the tract. If both vowels had a specific steady-state formant pattern, the diphthong was divided into two VtoV intervals; otherwise, it was treated as a single interval. For each utterance the average values of VtoV and %V were calculated. Figure 3 represents on the *x*-axis the percentage of vocalic portion in the utterance and on the *y*-axis the average values of VtoV.

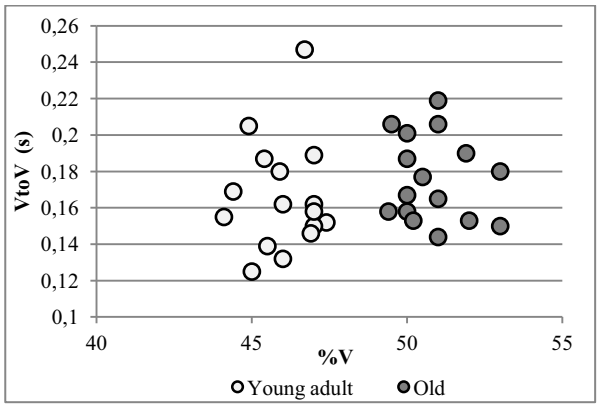

Figure 3: *%V and VtoV of Young Adult and Old speakers across all rates*.

The data plotted in figure 3 show that the utterance composition in terms of vocalic and consonantal portions is statistically different between young adults and old speakers ($p<0.01$). The former present %V values ranging from 44.1% to 47.4%. On the contrary, the performances of old speakers are characterized by higher %V values ranging from 49.4 to 53%. The intra-group differences both for young adults and old speakers are not statistically significant ($p>0.05$). These findings validate the data resulting from PA corpus and confirm the strict relationship between speaker age and vowel percentage in the utterance.

Chronological age being equal, the gender does not affect the %V values. In fact, in the same age-group, the differences between male and female voices are not statistically significant ($p>0.05$).

As for VtoV values, the data of both groups show considerable variations on the *y*-axis, reflecting the different speech rates of the utterances. The higher the VtoV values, the longer the time intervals between consecutive vocalic segments and, thus, the slower the speech. By contrast, the lower the VtoV values, the closer the vowel onset points and, thus, the faster the speech.

The data regarding the influence of speech rate on %V variations are shown in figures 4 and 5. The average values of %V for slow, medium, fast and very fast speech range from 45.5 to 46.5 in the group of Young Adults (fig. 4) and from 50 to 51.5 in the case of Old speakers (fig. 5). These slight variations, however, are not statistically significant for both age groups ($p>0.05$). These results confirm the outcomes of the study by Dellwo and Wagner [30] who found that %V remains rather constant across different speech rates.

The figure 6 plots the %V/VtoV values resulting from PA corpus and from the corpus of Young Adult and Old voices recorded at the four speech rates.

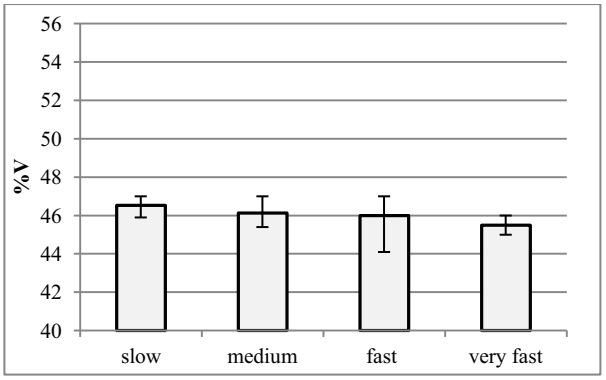

Figure 4: *Average %V values in Young Adults by speech rate*.

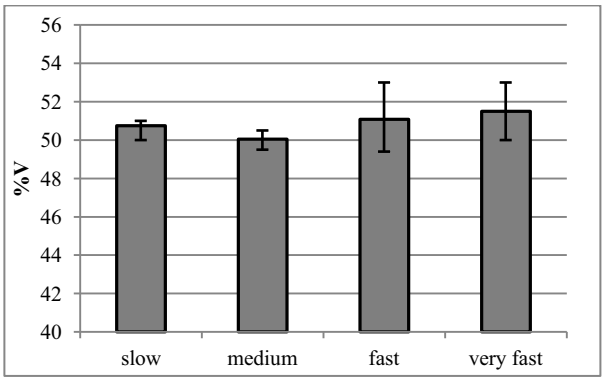

Figure 5: *Average %V values in Old Speakers by speech rate*.

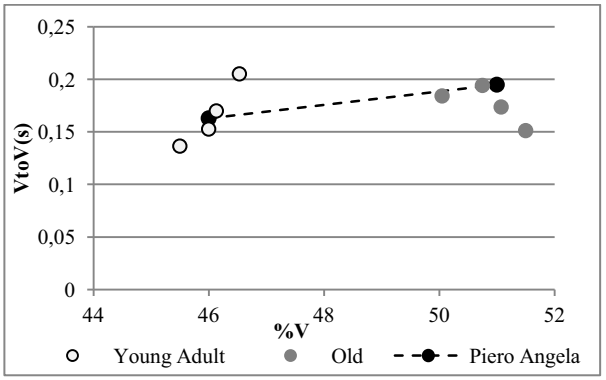

Figure 6: *Average %V/VtoV values of Young Adult and Old Speakers compared with PA corpus*.

As shown in figure 6, despite the marked differences between the two corpora (number and age distribution of the speakers, duration and speech type of the speech samples, corpus collection methods) the %V values of the 40-year old PA (46%) overlie upon those of Young Adults scored at medium and fast speech rate. Similarly, the %V values of 79-

year old PA lie in the range of the utterances produced by the old speakers at a low rate.

The data from PA speech provide important evidence on the relationship between speech rate and age that could not be otherwise inferred from the analysis of the corpus used in this study, as it was purposely elicited at four given speech rates. Communicative situation and speaker being equal, advancing age determines in PA both an increase in the %V and a decrease in speech rate: from medium to slow. This means that the slowing down of speech is due more to vowel lengthening rather than consonantal reduction.

On the contrary, chronological age being equal, speech style variations do not change %V because the speech deceleration/acceleration involves uniformly the whole utterance, at both level of consonantal and vocalic segments. Furthermore, the comparison between the two corpora suggests that in experimental conditions that do not impose explicit constraints on speech rate, 20-25 year old subjects would speak at medium rate and old speakers at slow rate.

4. Conclusions

Preliminary studies conducted on the speech produced by the same Italian speaker at the age of 40 and 79 have shown that, with advancing age, the %V increases, shifting towards the values of isomoraic languages. In order to verify whether variations in speech rhythm and %V were speaker- or age-dependent, four young adult and four old speakers of Italian were instructed to read four sentences of different length at four given speech rates. The results of the acoustic analysis of the corpus have confirmed the increase of %V in the aged voice, regardless of the speech rate. In Young Adults, aged between 20 and 25, %V ranges from 45.5 to 46.5, whereas in the elderly the same index increases significantly, amounting from 50 to 51.5. These data confirm that with advancing age the rhythm of Italian tends to move towards the isomoraic pattern.

In further steps, the research will consider different kinds of speech (informal and spontaneous) in order to better test the effect of aging on speech rate variation. Additionally, speakers belonging to the age bands not considered so far (25-75 years old) will be involved. This would allow to exactly define the temporal limits of %V variations and to shed light on the changes in the speech production mechanism responsible for these variations. Finally, the research could be also extended to speakers of other isosyllabic languages and of stress- and mora-timed languages.

5. References

[1] Eppley, B. D., and Mueller, P. B., "Chronological Age Judgments of Elderly Speakers: The Effects of Listeners' Age", Contemporary Issues in Communication Science and Disorders 28:5-8, 2001.

[2] Pear, T. H., Voice and Personality, London: Chapman & Hall, 1931.

[3] Ptacek, P. H., and Sander, E. K., "Age recognition from voice", Journal of Speech and Hearing Research, 9:273–277, 1966.

[4] Shipp, T., and Hollien, H., "Perception of the aging male voice", Journal of Speech and Hearing Research, 12:703–710, 1969.

[5] Pettorino, M., and Giannini, A., "The speaker's Age: A perceptual study", Proceedings of ICPhS XVII, Hong Kong, 1582-1585, 2011.

[6] Krauss, R. M., Freyberg, R., and Morsella, E., "Inferring speakers' physical attributes from their voices", Journal of Experimental Social Psychology, 38:618-625, 2002.

[7] Awan, S. N., "The aging female voice: acoustic and respiratory data", Clinical Linguistics & Phonetics, 20(2-3):171-180, 2006.

[8] Linville, S. E., "The aging voice", The American Speech-Language-Hearing Association (ASHA) Leader:12-21, 2004.

[9] Tompkins, C. A., Scharp, V. L., and Meigh, K. M., "Communication Disorders" in R. Schultz [Ed], The Encyclopedia of Aging, Fourth Edition, Vol. I A-K, 234-242, Springer Publishing, 2006.

[10] Wohlert, A., and Smith, A., "Spatiotemporal stability of lip movement in older adult speaker", Journal of Speech, Language and Hearing Research, 41:41-50, 1998.

[11] Scholtz, S., "Acoustic Analysis of Adult Speaker Age", in C. Müller [Ed], Speaker Classification I, LNAI 4343, 88-107, Springer, 2007.

[12] Linville, S. E., Vocal Aging, Singular, 2001.

[13] Linville, S. E., "Acoustic-perceptual studies of aging voice in women", Journal of Voice, 1:44-48, 1987.

[14] Jacques, R., and Rastatter, M., "Recognition of speaker age from selected acoustic features as perceived by normal young and older listeners", Folia Phoniatrica, 42:118–124, 1990.

[15] Traunmüller, H., and van Bezooijen, R., "The auditory perception of children's age and sex", ICSLP, 1171-1174, 1994.

[16] Ramig, L. A. and Ringel, R. L., "Effects of physiological aging on selected acoustic characteristics of voice", Journal of Speech and Hearing Research, 26: 22-30, 1983.

[17] Dehqan, A., Scherer, R. C., Dashti, G., Ansari-Moghaddam, A., and Fanaie, S., "The effects of aging on acoustic parameters of voice", Folia Phoniatrica et Logopaedica, 64: 265-270, 2012.

[18] Amerman, J. D., and Parnell, M. M., "Speech timing strategies in elderly adults", Journal of Phonetics, 20:65-76, 1992.

[19] Pettorino, M., Pellegrino, E., and Maffia, M. (forthcoming), " "Young" and "Old" Voice. The prosodic auto-transplantation technique for speaker's age recognition", Proceedings of 7th Speech Prosody, Dublin, 2014.

[20] Mehler, J., Dupoux, E., Nazzi T. & Dehaene-Lambertz, G., "Coping with linguistic diversity: The infant's viewpoint", in J. l. Morgan and K.E. Demuth [Eds.], Signal to Syntax: Bootstrapping from Speech to Grammar in early Acquisition 101–116, Erlbaum, 1996.

[21] Ramus, F., Nespor, M., and Mehler, J. "Correlates of linguistic rhythm in the speech signal", Cognition 73:265–292, 1999.

[22] Nespor, M., Shukla, M., and Mehler, J. "49. Stress-timed vs Syllable-timed languages" 1147-1159 http://www.sissa.it/cns/Articles/PhonCompanionTBC_049nespor%20et%20al.pdf. 2010, accessed in March 2014.

[23] Pettorino, M., Maffia, M., Pellegrino, E., Vitale, M. and De Meo, A., "VtoV: A perceptual cue for rhythm identification", in P. Mertens and I.C. Simon [Eds.], Proceedings of the Prosody-Discourse Interface Conference, 101-106, 2013.

[24] Tsujimura, N., [Ed], Japanese Linguistics: Critical Concepts. Volume I: Phonology and Morphology. Routledge Library of Modern Japan, 2005.

[25] Giannini, A., "Range di variabilità della velocità di articolazione in italiano", in M. Refice and M. Savino, M. [Eds], Atti del 28° Convegno Nazionale AIA Trani, 253-256, Addante, 2000.

[26] Magno Caldognetto, E., and Vagges, K., "Indici di fluenza, tipologia e distribuzione delle sillabe nel parlato spontaneo", Atti del XIX Convegno Nazionale A.I.A., Napoli, 423-429, 1991.

[27] Salza, P. L., "Velocità di eloquio e durata segmentale all'interno di parola: indagine preliminare", Atti del 19° Convegno Nazionale A.I.A., Napoli, 459-463, 1991.

[28] Zmarich, C., Magno Caldognetto, E., and Ferrero, F., "Analisi confrontativa di parlato spontaneo e letto: fenomeni macroprosodici e indici di fluenza", in Atti delle VII Giornate di Studio del GFS, Napoli, 111-139, Esagrafica, 1997.

[29] Boersma, P., "Praat, a system for doing phonetics by computer", *Glot International* 5(9-10): 341-345, 2001.

[30] Dellwo, V., and Wagner, P., "Relations between language rhythm and speech rate", in Proceedings of 15th ICPhS Barcelona, 471-474, 2003.