



# The Effect of Ageing on Speech Rhythm: A Study on Zurich German

*Elisa Pellegrino, Lei He, Volker Dellwo*

Department of Computational Linguistics, University of Zurich, Switzerland

{elisa.pellegrino, lei.he, volker.dellwo}@uzh.ch

## Abstract

Speech segmental and suprasegmental characteristics vary considerably across the life span, for example, due to degenerative changes in speech production mechanisms and neuro-muscular control. A great deal of research on the acoustic correlates of adult speakers' voice has focussed on changes in voice quality, vowel formant patterns,  $f_0$ , amplitude and speech rate. Only little attention has been paid on speech rhythm variability due to advancing age.

Here we quantified age-related rhythmic variability in terms of the durational characteristics of consonantal and vocalic intervals (henceforth CV intervals). We compared the segmental durational variability of two groups of Zurich German speakers. Group 1: 16 young adults, aged from 18 to 32 years; group 2: 10 older adults, aged from 66 to 81 years. For both groups we analyzed 20 sentences in Zurich German from the TEVOID Corpus. Between-speaker durational variability across age was quantified through a variety of interval-based metrics: segment rate, %V, deltaC, deltaV, VarcoC, VarcoV, rPVI-C and nPVI-V.

Results showed that rhythmic differences between younger and older adults are largely accountable for by speech rate differences. Segment rate, %V and raw measures of CV interval durational variability (deltaV, deltaC and rPVI-C) showed effects between younger and older adults. Rate normalized metrics (VarcoC, VarcoV and nPVI-V) did not differ significantly between the two age-groups.

**Index Terms:** speech rhythm, segmental durational variability, speech rate, ageing

## 1. Introduction

It is commonly known that age-related changes in speech production mechanisms as well as in neuro-muscular control determine considerable variation in speech production across the life span [1]. Stiffening of thorax, decreased lung capacity, and weakening of respiratory muscles are listed among the most typical changes in the respiratory system that have an impact on speech production. Ossification and calcification of laryngeal cartilages, thickening of laryngeal epithelium as well as the atrophy of vocal folds are, instead, some of the most salient changes occurring in the laryngeal system [1]-[5]. Atrophy of facial, mastication and pharyngeal muscles, decreased lip strength and reduced accuracy of the lower lip and jaw when performing rapid movements, are the most representative changes undergone by the supra-laryngeal system [1]-[5].

The process of ageing is also accompanied by changes in motor function. Neuropathy of nerve fibres, decline in the number of motor units and nerve cells, slowing down of conduction velocity, decline in dopamine level are only few

of the numerous changes occurring in central and peripheral nervous system that impact speech production [1].

Taken together, these changes alter considerably the way individuals speak over time both segmentally and suprasegmentally [4]-[7]. Although vocal ageing can be affected also by factors other than chronological age (i.e. environmental factors, lifestyle, smoking habits, medical conditions [8]), features that are typically associated with adult speakers' voice are: altered vowel formant frequency patterns, increased jitter and shimmer, increased breathiness, lengthening of vowels and stop consonants, reduced speech rate, decline in amplitude stability (at least in men), higher pitch voice in men and lower pitch voice in women [9]-[15].

How do these physiological and neurological changes across the life span affect the rhythm of speech? Cross-linguistic studies quantifying speech rhythm in terms of the durational characteristics of CV intervals have provided evidence that a) child speech rhythm presents higher proportion over which speech is vocalic, higher durational variability in consonantal intervals but lower variability in vocalic intervals compared to adult speech rhythm; b) over time child speech rhythm approximates to the adult profile [16],[17].

What has been investigated in far less detail, however, is the effect of healthy ageing on rhythmic characteristics of speech. Preliminary longitudinal and cross-sectional studies on Italian speakers support the view that speech rhythm characteristics vary as a function of age [18], [19]. [18] analyzed utterances of identical lexical content by one speaker at the age of 40 and 79. They found that %V as well as the mean duration of the interval between two consecutive vowel onset points (VtoV) increased significantly with advancing age. Based on these findings, [19] conducted a follow-up cross-sectional study to test whether the rhythmic variations found in [18] had been speaker- or age-dependent. In this study, 4 younger adults (aged between 20 and 25) and 4 older adults (aged between 75 and 80) were asked to read 4 sentences at 4 different articulation rates. The sentences varied from 20 to 35 syllables in length and each utterance was to be spoken in 5 sec. The results confirmed the increase of %V in the aged voices found in [18]. Regardless of speech rate differences, indeed, %V ranged from 45.5 to 46.5 for the group of younger adults, whereas for the older speakers the same index amounted from 50 to 51.5.

## 2. The study

### 2.1. Aim

The general aim of the present study was to investigate in depth whether degenerative changes in speech production mechanism as well as in neuro-muscular control have an

impact on the durational characteristics of consonantal and vocalic intervals. Compared to the research on age-related rhythmic variations mentioned above ([18], [19]) that:

- were based on the recordings of a small set of sentences spoken by a limited number of participants,
- quantified the differences between younger and older adults only in terms of %V and mean VtoV,

in the present study, we substantially increased the number of participants and sentences (see 2.2; 2.3), and investigated the durational variability across age through a variety of different interval-based measures (see 2.4).

## 2.2. Participants and Speech Material

We asked 10 Zurich German speakers, ranging in age from 66 to 81 years ( $M = 71.7$ ;  $SD = 4.9$ ), to read aloud a list of 256 sentences in Zurich German, from the TEVOID Corpus [20] (older adults, henceforth: OA). These sentences had been originally recorded by 16 speakers of Zurich German, aged between 18 and 32 years ( $M = 24.1$ ;  $SD = 3.8$ ) (younger adults, henceforth: YA).

Why did we collect read and not spontaneous speech? There is large evidence that rhythmic patterns are largely affected by the linguistic and prosodic structure of the utterance (syllabic composition, intonation, stress-pattern) and elicitation methods (sentences, reading a story, spontaneous speech) [21-23]. For this reason, it seemed that read speech possibly represented the kind of speech material most suitable for rhythmic analysis.

Given the evidence that rhythmic characteristics vary between Swiss German dialects [24], OA were selected based on strong screening criteria for Zurich German. Like YA, OA had been Zurich German monolinguals until 6 years old, had had Zurich German-speaking parents, were educated in Zurich and had lived in the city or in Zurich Oberland for at least five years before the recording sessions. YA and OA declared that they had no vision or hearing disabilities, nor recognized dyslexia. OA passed the Mini-Mental Status Examination (MMSE) [25].

OA and YA were recorded in the same sound treated booth at Zurich University. The TEVOID corpus produced by YA and OA will be hereafter called TEVOID\_YA and TEVOID\_OA respectively. Given that OA produced sentences with segmental mistakes and filled pauses, from the TEVOID\_OA we selected only the utterances that were produced with comparable segmental contents and without disfluencies across all speakers. For this reason, the final TEVOID corpus for OAs (henceforth called TEVOID\_OA\_FIN) consisted of 900 utterances (90 sentences \* 10 OA).

## 2.3. Annotation for rhythmic analysis

The first 20 utterances of TEVOID\_OA\_FIN were used for the present study. They were all declarative sentences, from 3 to 10 word-long, 19 main clauses and 1 complex clause (main clause and object clause separated by a comma).

All utterances were annotated in different tiers according to the segmentation criteria applied to the TEVOID\_YA [20]

- Tier 1 ('segments'), contained the manually annotated speech segments. Based on this tier, three other tiers containing consonantal and vocalic interval information were automatically derived.

- Tier 2 ('CV segments'), included information about whether the segment was consonantal or vocalic.
- Tier 3 ('CV segment intervals') and 4 ('CV intervals') segmented the signal in consecutive consonantal or vocalic intervals. Compared to tier 4, each interval of tier 3 contained also the number of consonantal or vocalic segments included in such interval.

For this study, the rhythmic analyses were run on 520 utterances, of which:

- 320 were extracted from TEVOID\_YA (20 sentences \* 16 YAs) and
- 200 were extracted from TEVOID\_OLD\_DEF (20 sentences \* 10 OAs).

For both corpora, we used the information contained in tier 1 to automatically calculate segment rate (segment/sec.) and in tier 4 to automatically compute the acoustic measures of speech rhythm, described in 2.4.

## 2.4. Acoustic measures of speech rhythm

In line with studies on adult and child speech rhythm [16] [17], we calculated:

- %V: the proportion over which speech is vocalic (%V) [26];
- deltaC and deltaV: standard deviation of the duration of consonantal and vocalic intervals [26].
- rPVI-C: average differences between two consecutive consonantal intervals [27].
- nPVI-V: rate-normalized average differences between two adjacent vocalic intervals [27].
- VarcoC and VarcoV: rate normalized standard deviation of consonantal and vocalic intervals [28], [29].
- Segment rate (segment/second).

DeltaC, deltaV and r-PVI-C are typically referred to as "raw" measures", since they are not normalized for speech rate, while VarcoC, VarcoV and n-PVI-V as "rate-normalized" metrics.

## 2.5. Hypotheses

Based on the literature on the acoustic correlates of adult speaker's voice as well as on previous findings on age-related rhythmic variations in Italian, we would expect OA to display slower articulation rate and higher %V. As regards the effect of the ageing process on CV intervals durational variability, we formulate the following hypotheses:

Hypothesis 1: If the ageing process affects both speech rate and the relative duration of consonantal and vocalic intervals, both raw and rate-normalized duration-based metrics should show significant effect between YA and OA.

Hypothesis 2: If the effect of ageing is limited to speech rate, but does not extend to the timing characteristics of consonantal and vocalic intervals, then the two age-groups should perform differently on raw-measures (deltaC, deltaV and r-PVI-C) but similarly on rate-normalized metrics (VarcoC, VarcoV and n-PVI-V).

## 2.6. Data analysis and Results

To analyse which rhythmic measures differ most between YA and OA, we ran Mixed-Effect Models with rhythm measures as dependent variables and age group ('young' and 'old') as fixed factor. Speakers and sentences were entered as random factors.

As summarized in Table 1, there is a significant effect of age group for the following rhythm measures: Segment Rate, %V, as well as all raw measures of CV interval durational variability (deltaC, deltaV, rPVI-C). Compared to YA, OA speak at significantly slower rate (fig. 1), but paused similarly to them. Both groups, indeed, typically read sentences without silent pauses, except for the complex sentence for which OA always produced a silence between the main and the subordinate clauses. OA exhibited higher %V (fig. 2), higher durational variability of CV intervals both when the variability is calculated on the whole utterance (deltaC and deltaV - figs. 3-4) and when it is computed between two consecutive intervals (r-PVI-C - fig. 5). Conversely, no significant differences are found between the two age groups in rate normalized CV interval durational variability (VarcoC, VarcoV, n-PVI-V) (table 2)

Table 1: Results from mixed-effect models for segment rate and four rhythm measures

Rhythm measures	df	X <sup>2</sup>	p
Segment rate	1	22.379	0.001
%V	1	21.094	0.001
delta C	1	7.0111	0.01
delta V	1	31.286	0.001
r-PVI-C	1	6.832	0.01

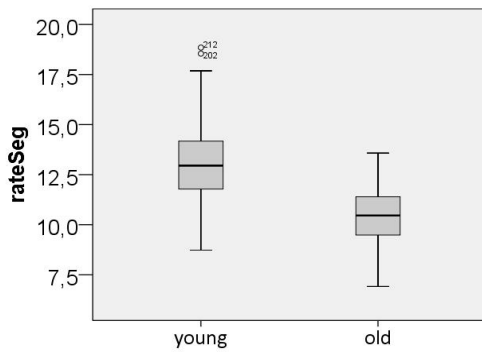


Figure 1: Boxplots of Segment Rate (segment/sec.) for Age Group.

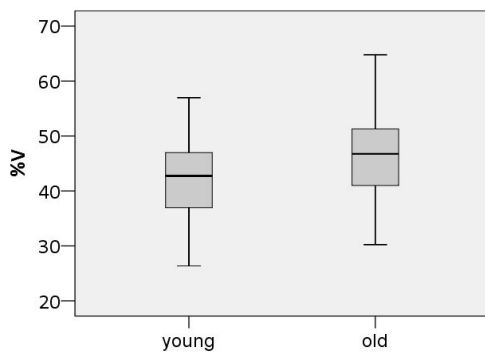


Figure 2: Box plot of %V for Age Group.

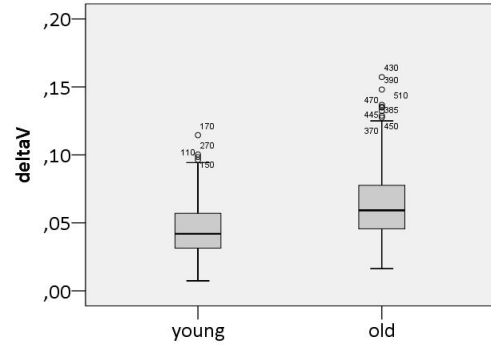


Figure 3: Box plot of vocalic durational variability in deltaV for Age Group.

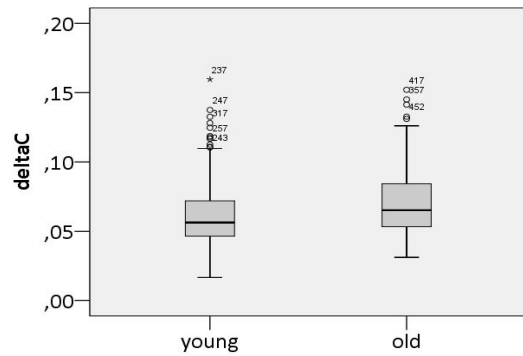


Figure 4: Box plot of consonantal durational variability in deltaC for Age Group.

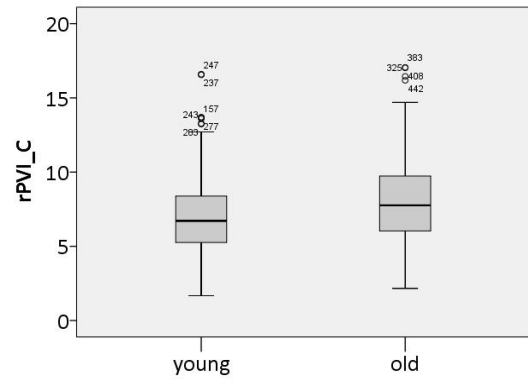


Figure 5: Box plot of consonantal durational variability in r-PVI-C for Age Group.

Table 2: Results from mixed-effect models for rate normalized rhythm measures.

Rhythm measures	df	X <sup>2</sup>	p
Varco C	1	0.6397	0.423
Varco V	1	1.9366	0.164
n-PVI-V	1	0.2136	0.643

### 3. Discussion

In line with previous findings on the acoustic correlates of adult speakers' voice, as well as on age-related rhythmic variations, the group of OA speaks more slowly and presents higher %V than YA. As regards the durational variability of CV intervals, the results of the present study point in favour of the hypothesis that healthy ageing selectively affects speech rate but not the relative duration of CV intervals (see 2.5). OA are higher compared to YA in terms of deltaV, deltaC and r-PVI-C. However, these metrics have been demonstrated to be affected by speech rate [28], [29]. On the other hand, OA approximate YA when compared in terms of rate-normalized measures (VarcoV, VarcoC, n-PVI-V).

What might determine the observed age-related changes? One first interpretation might lead to impute the decline in segment rate, the increase in vowel duration and in CV intervals durational variability to the generalized neuro-cognitive slowing [1] as well as to the degenerative changes in the supra-laryngeal systems [1] that alter the movements of the articulators.

On the other hand, findings from studies on the acoustic and articulatory variability across the age would support the view that the observed age-related rhythmic changes might be due to reasons other than diminished oro-facial strength or neuro-muscular degeneration [30]-[32]. [30], for example, found that older speakers with longer average vowel durations also tended to produce more acoustically distinct vowels. From an articulatory viewpoint, there is evidence that the slower speech is not imputable to constrained lip, jaw movements in older adults [31]-[32]. According to [32], indeed, older adults would reduce speech rate to compensate the decrease in lip and jaw stiffness when asked to produce speech at fast and very fast rates.

The reduced speech rate in OA can be possibly accounted for by their increasing difficulty of OA to connect the orthographic and the phonological units when reading aloud. In view of the transmission deficit theory, the lack of redundancy between these two nodes (phonological and orthographic) in a language like Swiss German could thus affect the phonological processing in later adulthood [33].

The discussion about whether the observed age-related variability is due to neuro-physiological changes or to a compensatory strategy to maintain the articulatory-acoustic precision is only of secondary importance to the present study. More relevant is the fact that by now there is evidence from datasets in Italian and Swiss German that healthy ageing affects speech rate but not the relative duration of CV intervals.

More research is, however, needed to better understand the role of healthy ageing in between-speaker durational variability. Clarifying this point might have a possible implication for the field of speech language pathology. There is evidence, indeed, that dysarthria determines considerable variation in the segmental durational variability between healthy and dysarthric patients [34], [35]. However, in [34], for example, the age range of both control and dysarthric speakers is rather wide. Knowing more about the effect of healthy ageing on between-speaker durational variability might contribute to disentangling the complex interplay of ageing and pathology in old dysarthric speakers

Last but not least, research on age-related rhythmic variations is also expected to ultimately contribute to forensic speech sciences. Numerous studies have demonstrated that:

a) the measures we investigated here can show robust variability between speakers [36-37] and

b) this variability is robust against several sources of within-speaker variability (i.e. speaking styles, speech rate and sentence structural characteristics [36-37].

However, given that in forensic casework, the time delay between the recording of a perpetrator and that of a suspect can sometimes be in the region of a few years, it is inevitable to understand more about the effects of age on such characteristics, if we want to apply time-domain measures of speaker individuality to forensic speaker comparison. For instance, [38] reports of some exemplar cases of mis-attribution of suspect utterances to the reference speaker, due to the misunderstanding of the relationship between age and the acoustic characteristic of voice.

### 4. Conclusions

The overall results of the present study show that the pattern of rhythmic variability between YA and OA is largely accountable for by speech rate differences. Segment rate, %V and measures of CV interval durational variability not normalized for speech rate (deltaV, deltaC and r-PVI-C) vary consistently between speakers as function of age group. The rate normalized metrics, on the other hand, are not crucially affected by the ageing process.

### 5. References

- [1] S. Scholtz, "Acoustic Analysis of Adult Speaker Age". In C. Müller [Ed.], *Speaker Classification I, LNAI 4343*, Berlin, Heidelberg: Springer, pp. 88-107, 2007.
- [2] S. N. Awan, "The aging female voice: acoustic and respiratory data", *Clinical Linguistics & Phonetics*, vol. 20, no 2-3, pp. 171-180, 2006.
- [3] S. E. Linville, "The aging voice", in *The American Speech Language-Hearing Association (ASHA) Leader*, pp. 12-21, 2004.
- [4] C. A. Tompkins, V. L. Scharp and K. M. Meigh, "Communication Disorders". In R. Schultz and L. Noelker [Eds.], *The Encyclopedia of Aging*, Fourth Edition, Vol. I A-L, New York: Springer Publishing, pp. 234-242, 2006.
- [5] A. Wohlert and A. Smith, "Spatiotemporal stability of lip movement in older adult speaker", *Journal of Speech, Language and Hearing Research*, vol. 41, pp. 41-50, 1998.
- [6] B. D. Eppley and P. B. Mueller, "Chronological Age Judgements of Elderly Speakers: The Effects of Listeners' Age", *Contemporary Issues in Communication Science and Disorders*, vol. 28, pp. 5-8, 2001.
- [7] T. Shipp and H. Hollien, "Perception of the aging male voice", *Journal of Speech and Hearing Research*, vol. 12, pp. 703-710, 1969.
- [8] S. E. Linville, Voice Disorders of Aging, in R. D. Kent [Ed.] *The MIT Encyclopedia of Communication Disorders*, Cambridge (MA): MIT Press, pp. 72-75, 2004.
- [9] S. E. Linville, *Vocal Aging*, San Diego: Singular Thomson Learning, 2001.
- [10] S. E. Linville, "Acoustic-perceptual studies of aging voice in women", *Journal of Voice*, vol. 1, pp. 44-48, 1987.
- [11] R. Jacques and M. Rastatter, "Recognition of speaker age from selected acoustic features as perceived by normal young and older listeners", *Folia Phoniatrica*, vol. 42, pp. 118-124, 1990.
- [12] H. Traummüller and R. van Bezooijen "The auditory perception of children's age and sex", in *ICSLP 94, Third International Conference on Spoken Language Processing (ICSLP 94)*,

- Yokohama, Japan, September 18-22, 1994, Proceedings, pp. 1171-1174, 1994.
- [13] L. A. Ramig and R. L. Ringel, "Effects of physiological aging on selected acoustic characteristics of voice", *Journal of Speech and Hearing Research*, vol. 26, pp. 22-30, 1983.
  - [14] A. Dehqan, R. C. Scherer, G. Dashti, Ansari-Moghaddam, A., and S. Fanaie, "The effects of aging on acoustic parameters of voice", *Folia Phoniatrica et Logopaedica*, vol. 64, pp. 265-270, 2012.
  - [15] J. D. Amerman and M. M. Parnell, "Speech timing strategies in elderly adults", *Journal of Phonetics*, vol. 20, pp. 65-76, 1992.
  - [16] E. Payne, B. Post, L. Astruc, P. Prieto, and M. d M Vanrell, "Measuring Child Rhythm", *Language and Speech*, pp. 1–27, 2011.
  - [17] E. Grabe, B. Post and I. Watson, "The acquisition of rhythmic patterns in English and French", in *ICPhS 1999, 14<sup>th</sup> International Congress of Phonetics Sciences, September 1-7, San Francisco, CA, USA, Proceedings*, pp. 1201–1204, 1999.
  - [18] M. Pettorino, E. Pellegrino and M. Maffia, " "Young" and "Old" Voice. The prosodic auto-transplantation technique for speaker's age recognition", in *Speech Prosody 7, 7<sup>th</sup> International Conference on Speech Prosody, May 20-23, Dublin, Ireland, Proceedings*, pp. 135-139, 2014.
  - [19] M. Pettorino and E. Pellegrino, "Age and Rhythmic Variations: A study on Italian", in *INTERSPEECH 2014, 15<sup>th</sup> Annual Conference of the International Speech Communication Association, September 14-18, Singapore, Proceedings*, pp. 1234-1237, 2014.
  - [20] V. Dellwo, A. Leeman and M. J. Kolly, "Speaker idiosyncratic rhythmic features in the speech signal", in *INTERSPEECH 2012, 13<sup>th</sup> Annual Conference of the International Speech Communication Association, September 9-13, Portland (OR), USA, Proceedings*, pp. 1584-1587, 2012.
  - [21] L. Wiget, L. White, B. Schuppler, L. Grenon, O. Rauch and S. L. Mattys, "How stable are acoustic metrics of contrastive speech rhythm", *The Journal of Acoustical Society of America*, vol. 127, no 3, 1559-1569, 2010.
  - [22] A. Arvaniti, "The usefulness of metrics in the quantification of speech rhythm", *Journal of Phonetics*, vol. 40, 351-373, 2012.
  - [23] P. Prieto, M. del Mar Vanrell, L. Astruc, E. Payne and B. Post, "Phonotactic and phrasal properties of speech rhythm. Evidence from Catalan, English, and Spanish", *Speech Commun.* vol. 54, pp. 681–702, 2012.
  - [24] A. Leemann, V. Dellwo, M. J. Kolly and S. Schmid, "Rhythmic variability in Swiss German dialects", in *Speech Prosody 6, 6<sup>th</sup> International Conference on Speech Prosody, May 22-25, Shanghai, China, Proceedings*, pp. 607-610, 2012.
  - [25] M. F. Folstein, S. E. Folstein and P. R. McHugh, "Mini-mental state". A practical method for grading the cognitive state of patients for the clinician, *Journal of Psychiatric Research*, vol. 12, no 3, pp. 189–198, 1975.
  - [26] F. Ramus, M. Nespore and J. Mehler, "Correlates of linguistic rhythm in the speech signal", *Cognition*, vol. 73, pp. 265-292, 1999.
  - [27] E. Grabe and L. Low, "Durational Variability in Speech and the Rhythm Class Hypothesis". In C. Gussenhoven [Ed.], *Papers in Laboratory Phonology 7*, Berlin: Mouton deGruyter, pp. 515-546, 2002.
  - [28] Dellwo, V. "Rhythm and Speech Rate: A Variation Coefficient for deltaC". In Karnowski, P., Szegedi, I. (Eds.). *Language and language-processing*, Frankfurt am Main: Peter Lang, pp. 231-241, 2006.
  - [29] L. White and S. L. Mattys, "Calibrating rhythm: First language and second language studies", *Journal of Phonetics*, vol. 35, pp. 501-522, 2007.
  - [30] A. R. Fletcher and M. McAuliffe, "The relationship between speech segment duration and vowel centralization in a group of older speakers", *The Journal of the Acoustical Society of America*, vol. 138, pp. 2-32, 2015.
  - [31] M. Bilodeau-Mercure and P. Trembley, "Age Differences in Sequential Speech Production: Articulatory and Physiological Factors", *Journal of the American Geriatrics Society*, vol. 64, no 11: e177–e182, 2016.
  - [32] A. S. Mefferd and E. E. Corder, "Assessing Articulatory Speed Performance as a Potential Factor of Slowed Speech in Older Adults", *Journal of Speech, Language, and Hearing Research*, 57, pp. 347–360, 2014.
  - [33] D. M. Burke and M. A. Shafto, "Language and aging". In Craik F.I.M. & Salthouse T.A. (Eds.), *The handbook of aging and cognition*, New Jersey: Lawrence Erlbaum Associate, pp. 373-443, 2008.
  - [34] J. M. Liss, L. White, S. L. Mattys, K. Lansford, A. J. Lotto, S. M. Spitzer and J. N. Caviness, Quantifying Speech Rhythm Abnormalities in the Dysarthrias, *Journal of Speech Language and Hearing Research*, vol. 52, no 5, pp. 1334–1352, 2009.
  - [35] M. Pettorino, M. G. Busà and E. Pellegrino, "Speech Rhythm in Parkinson's Disease: A Study on Italian", in *INTERSPEECH 2016, 17<sup>th</sup> Annual Conference of the International Speech Communication Association, September 8-12, San Francisco (CA), USA; Proceedings*, pp. 1958-1961, 2016.
  - [36] A. Leemann, V. Dellwo and M.-J. Kolly, "Forensic speaker identification based on speech temporal features", *Forensic Science International*, 238, pp. 59-67, 2014.
  - [37] V. Dellwo, A. Leemann and M.-J. Kolly, "Rhythmic variability between speakers: Articulatory, prosodic and lexical factors", *The Journal of the Acoustical Society of America*, vol. 137, no 3, pp. 1513-1528, 2015.
  - [38] R. Rhodes, "Aging effects on voice features used in forensic speaker comparison", *International Journal of Speech Language and Law*, vol. 24, no 2, 177–199, 2017.