Eppur si muove: Formant Dynamics is Relevant for The Study of Speech Aging Effects

Luciana Albuquerque[1,2,3,4,a], Catarina Oliveira[1,5,b], António Teixeira [1,3,c], and Daniela Figueiredo[2,5,d]

[1] *Institute of Electronics and Informatics Engineering of Aveiro, University of Aveiro, Aveiro, Portugal*
[2] *Center for Health Technology and Services Research, University of Aveiro, Aveiro, Portugal*
[3] *Department of Electronics Telecommunications and Informatics, University of Aveiro, Aveiro, Portugal*
[4] *Department of Education and Psychology, University of Aveiro, Aveiro, Portugal*
[5] *School of Health Science, University of Aveiro, Aveiro, Portugal*

Keywords: Aging Speech, Vowels, Dynamic Formant Frequencies, European Portuguese.

Abstract: The evidence have shown that speech change with age and the automatic speech recognition systems needs adaptation to older voices. Most of the acoustic studies about the age effects on speech production have focused on static approaches to obtain the vowel formants. However, vowel formant dynamics may also be important to characterize vowel quality and the age related changes. In this position paper the authors argue for the need to increase the use of dynamic information in acoustic studies. Among the main arguments, we can state that: speech is inherently dynamic; dynamic vowel formants improve the classification of vowels and dialects and play an important role in vowel perception; nowadays better tools allow to go beyond analysis of snapshots.

1 INTRODUCTION

As speech is a physiological signal that provides information at multiple levels concerning the linguistic aspects (e.g. words, message, accent, language) as well as the paralinguistic characteristics (e.g. gender, age, emotional state) (Sadjadi et al., 2016; Yue et al., 2014; Qawaqneh et al., 2017), the human speech can be used as an important cue to represent the person's age (Yue et al., 2014; Schötz, 2006).

The age effects on speech production mechanism have a significant impact on the acoustic measurements of speakers' vocal output (Xue and Hao, 2003; Schötz, 2006; Braun and Friebis, 2009). Despite the wide range of speech acoustic measurements that could be affected by age, in this position paper we will focus on vowel acoustics, namely on the different approaches to study the vowel formant frequencies.

[a] https://orcid.org/0000-0003-1654-3272
[b] https://orcid.org/0000-0002-3389-3082
[c] https://orcid.org/0000-0002-7675-1236
[d] https://orcid.org/0000-0002-3160-7871

1.1 Background Information

Population aging, while due primarily to lower fertility, also reflects a human success story of increased longevity (He et al., 2016). According to the World Health Organization (2012a,b) the number of people aged over 65 is increasing and Portugal is one of the developed countries with the highest rate of older population. Between 1970 and 2018, the percentage of people aged 65 and over increased from 9.7% to 21.8% Statistics Portugal (2019, 2015). This age group may increase from 2.1 million to 2.8 million between 2015 and 2080 in Portugal.

However, increasing longevity lead to new challenges, such as a pressure on health care costs, achieving life expectancy in good health, living independently Makiyama and Hirano (2017); He et al. (2016).

Aging involves changes at physiological, cognitive, psychological and social levels. Physiological age-related changes take place in different tissues and organs, and the human speech production mechanism is no exception Makiyama and Hirano (2017); Braun and Friebis (2009). Not only do the cognitive skills which are required in the planning process change with age, but also do initiation, phonation and artic-

Albuquerque, L., Oliveira, C., Teixeira, A. and Figueiredo, D.
Eppur si muove: Formant Dynamics is Relevant for The Study of Speech Aging Effects.
DOI: 10.5220/0010320902760283
In *Proceedings of the 14th International Joint Conference on Biomedical Engineering Systems and Technologies (BIOSTEC 2021) - Volume 4: BIOSIGNALS*, pages 276-283
ISBN: 978-989-758-490-9

ulation Braun and Friebis (2009).

Automatic extraction of speaker dependent characteristics, from a short speech utterance, is a challenging task, and it has a wide range of applications (Barkana and Zhou, 2015; Sadjadi et al., 2016; Qawaqneh et al., 2017). The age and gender detection of a speaker is a rapidly emerging field of research due to the continually growing interest in applications of communication, human–computer interface and natural spoken-dialog systems Barkana and Zhou (2015).

Overall, due to the aging speech alterations, the current automatic speech recognition (ASR) systems still do not work well with older people's speech Hämäläinen et al. (2012); Pellegrini et al. (2013). Besides, older adults have more difficulties in the interaction with computers, which represents a barrier to their access to new technologies Hämäläinen et al. (2012).

Although previous studies have presented some progress in extract and select features that represent the speaker's characteristics uniquely, the classification of speaker's age still has a lot to improve Qawaqneh et al. (2017), which is essential for the development of ASR systems suitable for older voices Vipperla et al. (2010). Moreover, a deeper knowledge of how speech changes with age is essential for clinical assessment and treatment of different speech disorders (Sataloff et al., 1997; Johns III et al., 2011), and also to provide information for other fields (e.g., phonetics, speech science, forensic linguistics and biometric recognition) (Kent and Vorperian, 2018; Lanitis, 2010).

2 STATE OF THE ART: AGING ACOUSTIC CHANGES

During the natural process of aging anatomical and physiological changes in the speech organs occur, which are reflected in the variation of several acoustic parameters, more specifically in the decrease of the speaking rate, in the increase of speech pauses, in the variation of the fundamental frequency and in the changes in the pattern of the formant frequencies, among others Linville and Rens (2001); Schötz (2006).

The magnitude of the speech acoustic changes depends upon the individual, as the voice is intricately linked to the dynamics of the speech organs Makiyama and Hirano (2017). Furthermore there are substantial gender differences in the extent and timing of the aging process Linville (2001); Makiyama and Hirano (2017); Schötz (2006).

Numerous studies have focused on age-related changes on F1 and F2, neglecting higher formants, and using a static approach to obtain the formant frequencies.Moreover, the results across studies are highly inconsistent. There are studies that have shown an age-dependent formant frequency lowering Linville and Rens (2001); Xue and Hao (2003); Watson and Munson (2007); Harrington et al. (2007); Decoster and Debruyne (1999) and others that have reported no changes in formant frequencies Fletcher et al. (2015); Sebastian et al. (2012). In Schötz (2006) and Eichhorn et al. (2018) there are vowels that presented a different pattern of formant frequencies variation with age and gender. In addition, some studies have referred a centralization of the vowel space in older speakers, which should result in a movement to the centroid of formant space Schötz (2006); Rastatter and Jacques (1990); Rastatter et al. (1997); Torre III and Barlow (2009); Mertens et al. (2020).

Concerning temporal information, it has often been noted that older adults use slower speaking rates (Linville, 2001; Schötz, 2006) and therefore produce vowels with longer durations Linville (2001); Schötz (2006); Fletcher et al. (2015); Benjamin (1982); Fougeron et al. (2018).

To the best of our knowledge, there are a few data available about the age effects on dynamic formant frequencies. Only Jacewicz et al. (2011a), in a cross-generational and cross-dialectal study, revealed substantial differences both in formant dynamics and vowel dispersion in the acoustic space as a function of age group.

2.1 European Portuguese Vowels

Most previous studies on European Portuguese (EP) vowels rely on single acoustic measurements of formant frequencies at the middle of the vowel (Pellegrini et al., 2013; Albuquerque et al., 2014, 2019; Oliveira et al., 2012).

The few data available for the EP have shown that age changes in vowel formant frequencies are not consistent and seem to be different among vowels (Albuquerque et al., 2019, 2014). These age-related acoustic changes might occur due to specific articulatory adjustments of the older speakers during speech, rather than generalized processes such as lengthening of the vocal tract (Xue and Hao, 2003; Eichhorn et al., 2018). Additionally, Albuquerque et al. (2019) showed a trend towards the centralization of vowels' space with age, mainly for males. This might indicate that articulation capability of males deteriorate with aging (Arias-Vergara et al., 2017). Further, vowel duration has presented consistent results in EP, showing a significantly increase with aging in both genders (Pellegrini et al., 2013; Albuquerque et al., 2014, 2019).

As far as we know, only Albuquerque et al. (2020) indirectly studied the age effect in vowels' formants dynamics and showed that dynamic measurements of F1-F3 result in higher rates of age group classification (senior/non-senior).

3 STATUS QUO: STATIC IS MORE THAN ENOUGH

The vowel formant measurements have a long history in the study of speech production, mainly because formant descriptions are suited to articulatory interpretations of acoustic data and formant frequencies reflect the length and configuration of the vocal tract McDougall and Nolan (2007); Kent and Vorperian (2018)

Previous studies have identified the first two or three formant frequencies as crucial acoustic correlates for the identification of vowels Peterson and Barney (1952); Watson and Harrington (1999); Almurashi et al. (2019); Adank et al. (2004); Kent and Vorperian (2018); Themistocleous (2017). Still, vowel duration Almurashi et al. (2019); Albuquerque et al. (2020) and fundamental frequency Zahorian and Jagharghi (1993) as additional cues have been reported to play a role in vowel discrimination.

The static approach has dominated acoustic analysis of vowels and is a convenient simplification because the vowel's midpoint is the section of the vowel that is least influenced by the contextual effects Watson and Harrington (1999); Peterson and Barney (1952); Kent and Vorperian (2018); Hillenbrand (2013) and also corresponds to the target position (i.e., vowel target), where a minimal shift in formant frequencies is seen Harrington (2010); Almurashi et al. (2019); Van der Harst and Van de Velde (2014).

Static vowel features may be sufficient for vowel classification, as it has previously been shown for some languages Williams et al. (2015); Sarvasy et al. (2020); Almurashi et al. (2019).

Additionally, also cross-dialectal studies have reported that vowel steady states are major acoustic vowel features Ewald et al. (2017); Van der Harst and Van de Velde (2014).

3.1 Vowel Dynamics is a Great Challenge

Despite the fact that vowels show a significant amount of spectral movement throughout their course Hillenbrand (2013); Williams and Escudero (2014); Sarvasy et al. (2019), the analysis of the vowel dynamics is a great challenge. For that, a solution has been to make measurements at single time point Kent and Vorperian (2018). The data obtained with this static approach have the advantage of simplicity and economy Kent and Vorperian (2018).

In a dynamic approach to extract vowel formants, with a large number of time points, the formant measures can wander in seemingly erratic directions through the course of a vowel Thomas (2011). In addition, the plots to represent the vowels become more difficult to interpret, making the comparisons between vowels challenging Thomas (2011).

Additionally, more complex models that allow to analyze formant means, as well as the direction and magnitude of the formant change, saturate on the amount of speaker-specific information which can be extracted and could begin fitting noise Watson and Harrington (1999); Ewald et al. (2017); Williams and Escudero (2014).

4 ARGUMENTS FOR OUR THESIS: DYNAMICS HAS POTENTIAL

Having presented in the previous section our perception of the main counterclaims, we now present various arguments to support our position.

Although the effectiveness of the formant frequencies in vowel separation is indisputable, some studies have also recognized that temporal information has also been important for characterizing vowel quality Watson and Harrington (1999); Almurashi et al. (2019); Williams and Escudero (2014). However, this does not imply that duration is the best or even the most relevant way to discriminate vowels Sandoval and Utianski (2015). Whereas vowel duration is sensitive to speaking rate, formant trajectory computed relative to vowel duration is not Sandoval and Utianski (2015).

Therefore, dynamic approaches based on formant trajectories or combinations of measurement points that sample the vowel formant pattern should be taken into account Kent and Vorperian (2018).

Despite the fact that vowel formants may vary to some extent according to phonetic context, some formant movement may occur due to vowel inherent spectral change (VISC) Williams et al. (2015), which is defined by Nearey and Assmann (1986) as the "relatively slowly varying changes in formant frequencies associated with vowels themselves, even in the absence of consonantal context".

4.1 Speech is Dynamic and is Affected by Age

In speech production the jaw, tongue and other articulators move continuously in space and time through several articulatory postures per second Rogers et al. (2013). As a result, speech is inherently dynamic Yuan (2013).

Dynamic features of speech offer greater scope for variation among speakers, as they reflect the movement of the individual's speech organs, as well as anatomical dimensions McDougall and Nolan (2007). The explanations that have been advanced to account for age-related changes in vowel formant frequencies have referred alterations in both dimensions, mentioned above. Besides, for EP vowels, Albuquerque et al. (2019) indicated that older speakers might present specific articulatory adjustments during speech. Therefore, a dynamic approach to studying the age effect on vowel production might be important, since a static approach mainly demonstrates anatomical differences among speakers.

Furthermore, the static measures do not address information about how formant frequency changes in time (Fox and Jacewicz, 2009).

4.2 Better Tools to go Beyond Snapshots Analysis

For many years the methods and tools available for researchers made only viable a small amount of measures Van der Harst and Van de Velde (2014), but all changed with the availability of programmable tools such as Praat, making the extraction of several measures over time as easy as extraction just one.

Meanwhile two sets of approaches to study the dynamic properties of vowels have been developed: a series of successive time points (multiple time point approach) and by curve-fitting Van der Harst and Van de Velde (2014).

In the multiple time point approach, measures may involve comparing formant frequencies from two or more discrete time points during vowel duration (e.g., two points Morrison (2013); Adank et al. (2004); three points (i.e., the three-point model) Hillenbrand et al. (1995); Almurashi et al. (2019)); five points Fox and Jacewicz (2009); or thirty points Williams and Escudero (2014)).

Alternatively, dynamic variations in the formants F1 and F2 may be characterized by different measures such as trajectory length (TL) and the spectral roc Fox and Jacewicz (2009). Even though these latter metrics incorporate more detailed spectral information, they do not account for the directionality of the change (i.e., if the frequencies actually increase or decrease over time) Williams and Escudero (2014).

On the other hand, vowel dynamics can also be expressed by curve-fitting parameterizations, by fitting parametric curves such as polynomials Themistocleous (2017) or discrete cosine transforms (DCTs) Elvin et al. (2016); Williams and Escudero (2014); Sarvasy et al. (2020); Watson and Harrington (1999) to formant contours for quantifying the shape of complex curve Brandt et al. (2018); Van der Harst and Van de Velde (2014). These approaches allow to analyze formant means as well as the direction and magnitude of the formant change Ewald et al. (2017); Williams and Escudero (2014). Furthermore, these avoid an arbitrary choice of one or more vowel targets, which can be tricky when a vowel appears either not having a steady-state section, or if the formants reach a minima or maxima at different times (Watson and Harrington, 1999).

Additionally, namely for recordings with poor quality, formant analysis using averaged values over multiple time points could be more reliable than only using a single time point measurement Sarvasy et al. (2020).

4.3 Dynamic Improve Classification Performance

A number of studies have reported that incorporating measures of formant dynamics enhances the classification of vowels based on acoustic Almurashi et al. (2019); Elvin et al. (2016); Yuan (2013); Jacewicz et al. (2011b); Jacewicz and Fox (2013); Jacewicz et al. (2009); Williams and Escudero (2014); Al-Tamimi (2007); Chittaragi and Koolagudi (2019); Zahorian and Jagharghi (1993). For instance, on Hijazi Arabic (HA) vowels, Almurashi et al. (2019) revealed that the static approach was sufficient for vowel classification, but multiple time point approaches performed better than a static approach Almurashi et al. (2019). Nonetheless, dynamic acoustic properties for vowel classification has not presented the same importance for all languages Sarvasy et al. (2020).

Williams and Escudero (2014); Elvin et al. (2016) agree that in addition to formant trajectory means, duration, magnitude and direction of formant trajectory slope are essential acoustic parameters for representing the English vowels. Also, they concluded that formant curvature (represented by the second DCT coefficients) was not necessary for classifying vowels (Elvin et al., 2016; Williams and Escudero, 2014), but can aid with more fine-grained/ subtle phonetic information from different speakers or different dialects (Elvin et al., 2016).

Furthermore, formant dynamics plays a major role in determining cross-dialectal acoustic differences for some vowels Van der Harst and Van de Velde (2014); Jacewicz and Fox (2013); Williams and Escudero (2014). In Van der Harst and Van de Velde (2014) both the multiple time point approach and the curve-fitting parameterization proved to be a clear improvement on the static approach to describe regional variation in the dynamics of vowel formants of Standard Dutch.

Formant dynamics are also useful for improving the within-class separation of the Australian English tense vowels from their lax counterparts Watson and Harrington (1999).

In addition, spectral change patterns may provide vowel phonetic details that are relevant in second-language (L2) learning (e.g., Jin and Liu (2013)) and, therefore, may prove to be useful for predicting L2 difficulties Elvin et al. (2016).

Although an approach only with two time points was performed equally well in distinguishing vowel categories as more sophisticated parametric curve approaches Zahorian and Jagharghi (1993); Morrison (2013), a whole trajectory approach based on parametric curves outperforms a two time points approach for extract speaker information Morrison (2013). Therefore, the measurements at more time points pay off in return to more social information than static measures. Van der Harst and Van de Velde (2014)

4.4 Dynamics is Relevant to Vowel Perception

The dynamics of formant-frequency patterns has been reported to play an important role in vowel perception, mainly for English (Hillenbrand, 2013; Jin and Liu, 2013; Jacewicz and Fox, 2012; Nearey and Assmann, 1986; Strange et al., 1983; Strange, 1989).

Perceptual studies have shown that: vowel steady-states can be removed with little or no effect on vowel intelligibility; vowels with stationary spectral patterns are not well identified; and also vowels in consonant context are more accurately identified than isolated vowels (Strange et al., 1983; Strange and Bohn, 1998; Hillenbrand, 2013).

Therefore, the dynamic vowel formants provide essential information about the characteristic of the vowels (Strange et al., 1983), which support a dynamic specification of vowel theory over static target theories (Strange and Jenkins, 2013). However, Jacewicz and Fox (2012) demonstrated that formant dynamics (i.e., VISC) can play a significant role in error identification of some vowels more than others by listeners.

5 CONCLUSION

The authors strongly believe in the potential of dynamic measurements of vowel formants for vowel classification. Also that these dynamic measurements provide useful information about the speakers identity. As the vowel dynamics could be affected by age, this information may be important to improve automatic extraction of speaker age.

Summing up, the use of dynamic approaches as a proxy for kinematic movement may be useful as a means to track changes of the normal speech with age. This could be validated further in experiments to determine how formant trajectories change with age for each gender.

As the population aging increases, the world is facing new challenges He et al. (2016), and the automatic extraction from speech of speaker dependent characteristics, such as age, has a wide range of applications that could be useful to improve the quality of life of older people Sadjadi et al. (2016); Yue et al. (2014).

6 FUTURE DIRECTIONS: A PROPOSAL

In order to explore the impact of dynamic vowel formants in EP vowel classification throughout life span, there is still a lot of work to be done, and the following actions are required:

- to apply different dynamic approaches to existing speech corpora of EP (e.g., Hämäläinen et al. (2012); Albuquerque et al. (2019)) in order to obtain the dynamic vowel formants (i.e., F1, F2, and F3) and to investigate which ones fit better the acoustic changes with age;
- to apply several classification algorithms (e.g., Support Vector Machine (SVM) and (Deep) Neural Networks) to the same body of acoustic vowel data, to investigate the performances of static and dynamic information in vowel and age classification tasks in many ways:
  - to explore the classification of the EP vowel according to their static and dynamic properties;
  - to analyze percentage of errors in vowel classification of the EP vowels by age, in order to

analyze if there is an age group and/or vowels more effectively classified based on dynamic information compared with the static alone;

- to examine the age classification performances based on static information alone and on both static and dynamic information of vowel formants;
- to identify which are the dynamic cues with more impact on age and vowel classification performance.

In our understanding, future work should seek to determine the extent to which the dynamic aspects of the vowel's acoustic signal contribute to its identification beyond the static information which is available at the vowel target, and whether these dynamic aspects of the vowels change with age. Also if there are observable differences between genders.

ACKNOWLEDGEMENTS

This research was financially supported by the project Vox Senes POCI-01-0145-FEDER-03082 (funded by FEDER, through COMPETE2020 - Programa Operacional Competitividade e Internacionalização (POCI), and by national funds (OE), through FCT/MCTES), by the grant SFRH/BD/115381/2016 and by IEETA (UIDB/00127/2020).

REFERENCES

Adank, P., Van Hout, R., and Smits, R. (2004). An acoustic description of the vowels of Northern and Southern Standard Dutch. *J. Acoust. Soc. of Am.*, 116(3):1729–1738.

Al-Tamimi, J. (2007). Static and dynamic cues in vowel production: A cross dialectal study in Jordanian and Moroccan Arabic. In *16th International Congress of Phonetic Sciences (ICPhS)*, pages 541–544, Saarbrücken, Germany.

Albuquerque, L., Oliveira, C., Teixeira, A., Sa-Couto, P., and Figueiredo, D. (2019). Age-related changes in European Portuguese vowel acoustics. In *INTERSPEECH*, pages 3965–3969, Graz, Austria.

Albuquerque, L., Oliveira, C., Teixeira, A., Sa-Couto, P., Freitas, J., and Dias, M. S. M. (2014). Impact of age in the production of European Portuguese vowels. In *INTERSPEECH*, pages 940–944, Singapore.

Albuquerque, L., Teixeira, A., Oliveira, C., and Figueiredo, D. (2020). The effect of dynamic acoustic cues on age classification. In *SPPL2020: 2nd Workshop on Speech Perception and Production across the Lifespan (Poster)*, page 81.

Almurashi, W., Al-Tamimi, J., and Khattab, G. (2019). Static and dynamic cues in vowel production in Hijazi Arabic. In *19th ICPhS*, pages 3468–3472, Newcastle.

Arias-Vergara, T., Vásquez-Correa, J. C., and Orozco-Arroyave, J. R. (2017). Parkinson's disease and aging: analysis of their effect in phonation and articulation of speech. *Cognitive Computation*, 9(6):731–748.

Barkana, B. D. and Zhou, J. (2015). A new pitch-range based feature set for a speaker's age and gender classification. *Applied Acoustics*, 98:52–61.

Benjamin, B. J. (1982). Phonological performance in gerontological speech. *Journal of Psycholinguistic Research*, 11(2):159–167.

Brandt, E., Zimmerer, F., Andreeva, B., and Möbius, B. (2018). Impact of prosodic structure and information density on dynamic formant trajectories in German. In *9th International Conference on Speech Prosody*, pages 119–123.

Braun, A. and Friebis, S. (2009). Phonetic cues to speaker age: A longitudinal study. In Grewendorf, G. and Rathert, M., editors, *Formal Linguistics and Law*, pages 141–162. De Gruyter Mouton, Berlin.

Chittaragi, N. B. and Koolagudi, S. G. (2019). Acoustic-phonetic feature based Kannada dialect identification from vowel sounds. *International Journal of Speech Technology*, 22(3):1099–1113.

Decoster, F. and Debruyne, W. (1999). Acoustic differences between sustained vowels perceived as young or old. *Log Phon Vocol*, 24(1):1–5.

Eichhorn, J. T., Kent, R. D., Austin, D., and Vorperian, H. K. (2018). Effects of Aging on Vocal Fundamental Frequency and Vowel Formants in Men and Women. *Journal of Voice*, 32(5):644.e1–644.e9.

Elvin, J., Williams, D., and Escudero, P. (2016). Dynamic acoustic properties of monophthongs and diphthongs in Western Sydney Australian English. *J. Acoust. Soc. Am.*, 140(1):576–581.

Ewald, O., Liina Asu, E., and Schötz, S. (2017). The formant dynamics of long close vowels in three varieties of Swedish. In *INTERSPEECH*, pages 1412–1416, Stockholm, Sweden. ISCA.

Fletcher, A. R., McAuliffe, M. J., Lansford, K. L., and Liss, J. M. (2015). The relationship between speech segment duration and vowel centralization in a group of older speakers. *J. Acoust. Soc. Am.*, 138(4):2132–2139.

Fougeron, C., D'Alessandro, D., and Lancia, L. (2018). Reduced coarticulation and aging. *J. Acoust. Soc. Am.*, 144(3):1905.

Fox, R. A. and Jacewicz, E. (2009). Cross-dialectal variation in formant dynamics of American English vowels. *J. Acoust. Soc. Am.*, 126(5):2603–2618.

Hämäläinen, A., Pinto, F. M., Dias, M. S., Júdice, A., Pires, C. G., Teixeira, V. D., Calado, A., and Braga, D. (2012). The First European Portuguese Elderly Speech Corpus. In *IberSPEECH 2012: "VII Jornadas en Tecnología del Habla" and "III Iberian SLTech"*, Madrid, Spain.

Harrington, J. (2010). *Phonetic analysis of speech corpora*. John Wiley & Sons.

Harrington, J., Palethorpe, S., and Watson, C. I. (2007). Age-related changes in fundamental frequency and formants: a longitudinal study of four speakers. In *INTERSPEECH*, pages 2753–2756, Belgium.

He, W., Goodkind, D., and Kowal, P. R. (2016). An aging world: 2015. *International Population Reports*, P95/16-1.

Hillenbrand, J., Getty, L. A., Clark, M., and Wheeler, K. (1995). Acoustic characteristics of American English vowels. *J. Acoust. Soc. of Am.*, 97(5 Pt 1):3099–3111.

Hillenbrand, J. M. (2013). Static and dynamic approaches to vowel perception. In Morrison, G. S. and Assmann, P. F., editors, *Vowel inherent spectral change*, pages 9–30. Springer.

Jacewicz, E. and Fox, R. A. (2012). The effects of cross-generational and cross-dialectal variation on vowel identification and classification. *J. Acoust. Soc. Am.*, 131(2):1413–1433.

Jacewicz, E. and Fox, R. A. (2013). Cross-dialectal differences in dynamic formant patterns in American English vowels. In *Vowel inherent spectral change*, pages 177–198. Springer.

Jacewicz, E., Fox, R. A., O'Neill, C., and Salmons, J. (2009). Articulation rate across dialect, age, and gender. *Language variation and change*, 21(02):233–256.

Jacewicz, E., Fox, R. A., and Salmons, J. (2011a). Cross-generational vowel change in American English. *Language variation and change*, 23(1):45–86.

Jacewicz, E., Fox, R. A., and Salmons, J. (2011b). Vowel change across three age groups of speakers in three regional varieties of American English. *Journal of Phonetics*, 39(4):683–693.

Jin, S.-H. and Liu, C. (2013). The vowel inherent spectral change of English vowels spoken by native and non-native speakers. *J. Acoust. Soc. Am.*, 133(5):EL363–EL369.

Johns III, M. M., Arviso, L. C., and Ramadan, F. (2011). Challenges and opportunities in the management of the aging voice. *Otolaryngology - Head and Neck Surgery*, 145(1):1–6.

Kent, R. D. and Vorperian, H. K. (2018). Static Measurements of Vowel Formant Frequencies and Bandwidths: A Review. *Journal of Communication Disorders*, 74:74–97.

Lanitis, A. (2010). A survey of the effects of aging on biometric identity verification. *International Journal of Biometrics*, 2(1):34–52.

Linville, S. E. (2001). *Vocal aging*. Singular Thomson Learning, Australia, San Diego.

Linville, S. E. and Rens, J. (2001). Vocal Tract Resonance Analysis of Aging Voice Using Long-Term Average Spectra. *Journal of Voice*, 15(3):323–330.

Makiyama, K. and Hirano, S. (2017). Aging Voice.

McDougall, K. and Nolan, F. (2007). Discrimination of speakers using the formant dynamics of/u:/in British English. In *International Congress of Phonetic Sciences (ICPhS XVI)*, pages 1825–1828, Saarbrücken, Germany.

Mertens, J., Mücke, D., and Hermes, A. (2020). Aging effects on prosodic marking in German: An acoustic analysis. In *2nd Workshop on Speech Perception and Production across the Lifespan (Poster)*, London. UCL.

Morrison, G. S. (2013). Vowel inherent spectral change in forensic voice comparison. In Morrison, G. S. and Assmann, P. F., editors, *Vowel Inherent Spectral Change*, pages 263–282. Springer Berlin Heidelberg.

Nearey, T. M. and Assmann, P. F. (1986). Modeling the role of inherent spectral change in vowel identification. *J. Acoust. Soc. Am.*, 80(5):1297–1308.

Oliveira, C., Cunha, M. M., Silva, S., Teixeira, A., and Sa-Couto, P. (2012). Acoustic analysis of European Portuguese oral vowels produced by children. In *IberSPEECH*, volume 328, pages 129–138, Madrid, Spain.

Pellegrini, T., Hämäläinen, A., de Mareüil, P. B., Tjalve, M., Trancoso, I., Candeias, S., Dias, M. S., and Braga, D. (2013). A corpus-based study of elderly and young speakers of European Portuguese: acoustic correlates and their impact on speech recognition performance. In *INTERSPEECH*, pages 852–856, Lyon.

Peterson, G. E. and Barney, H. L. (1952). Control methods used in a study of the vowels. *J. Acoust. Soc. of Am.*, 24:175.

Qawaqneh, Z., Mallouh, A. A., and Barkana, B. D. (2017). Deep neural network framework and transformed MFCCs for speaker's age and gender classification. *Knowledge-Based Systems*, 115:5–14.

Rastatter, M. P. and Jacques, R. D. (1990). Formant frequency structure of the aging male and female vocal tract. *Folia phoniatrica*, 42(6):312–319.

Rastatter, M. P., McGuire, R. A., Kalinowski, J., and Stuart, A. (1997). Formant frequency characteristics of elderly speakers in contextual speech. *Folia Phoniatrica et Logopaedica*, 49(1):1–8.

Rogers, C. L., Glasbrenner, M. M., DeMasi, T. M., and Bianchi, M. (2013). Vowel inherent spectral change and the second-language learner. In Morrison, G. S. and Assmann, P. F., editors, *Vowel Inherent Spectral Change*, pages 231–259. Springer Berlin Heidelberg.

Sadjadi, S. O., Ganapathy, S., and Pelecanos, J. W. (2016). Speaker age estimation on conversational telephone speech using senone posterior based i-vectors. In *International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pages 5040–5044. IEEE.

Sandoval, S. and Utianski, R. L. (2015). Average Formant Trajectories. *Preprint submitted to Journal ofPhonetics*.

Sarvasy, H., Elvin, J., Li, W., and Escudero, P. (2019). Vowel acoustic of Nungon, Papua New Guinea. In *ICPhS'19*, pages 1714–1718, Melbourne, Australia.

Sarvasy, H., Elvin, J., Li, W., and Escudero, P. (2020). An acoustic phonetic description of Nungon vowels. *J. Acoust. Soc. Am.*, 147(4):2891–2900.

Sataloff, R. T., Caputo Rosen, D., Hawkshaw, M., and Spiegel, J. R. (1997). The aging adult voice. *Journal of Voice*, 11(2):156–160.

Schötz, S. (2006). *Perception, analysis and synthesis of speaker age*, volume 47. Linguistics and Phonetics, Lund University.

Sebastian, S., Babu, S., Oommen, N. E., and Ballraj, A. (2012). Acoustic measurements of geriatric voice. *Journal of Laryngology and Voice*, 2(2):81–84.

Statistics Portugal (2015). Envelhecimento da população residente em Portugal e na União Europeia (Aging of the resident population in Portugal and the European Union). *Destaque: informação à comunicação social*.

Statistics Portugal (2019). Estimativas de População Residente em Portugal - 2018 (Estimates of resident population in Portugal - 2018). *Destaque: informação à comunicação social*.

Strange, W. (1989). Dynamic Specification of Coarticulated Vowels Spoken in Sentence Context. *J. Acoust. Soc. Am.*, 85(5):2135–2153.

Strange, W. and Bohn, O.-S. (1998). Dynamic specification of coarticulated German vowels: Perceptual and acoustical studies. *J. Acoust. Soc. Am.*, 104(1):488–504.

Strange, W. and Jenkins, J. J. (2013). Dynamic specification of coarticulated vowels. In Morrison, G. S. and Assmann, P. F., editors, *Vowel Inherent Spectral Change*, pages 87–115. Springer Berlin Heidelberg.

Strange, W., Jenkins, J. J., and Johnson, T. L. (1983). Dynamic specification of coarticulated vowels. *J. Acoust. Soc. Am.*, 74(3):695–705.

Themistocleous, C. (2017). Dialect classification using vowel acoustic parameters. *Speech Communication*, 92:13–22.

Thomas, E. (2011). *Sociophonetics: An Introduction*. Palgrave Macmillan.

Torre III, P. and Barlow, J. A. (2009). Age-related changes in acoustic characteristics of adult speech. *Journal of Communication Disorders*, 42:324–333.

Van der Harst, S. and Van de Velde, H. (2014). Variation in Standard Dutch vowels: The impact of formant measurement methods on identifying the speaker's regional origin. *Language variation and change*, 26(2):247–272.

Vipperla, R., Renals, S., and Frankel, J. (2010). Ageing voices: The effect of changes in voice parameters on ASR performance. *J. Aud. Speech Music Process*, pages 1–10.

Watson, C. I. and Harrington, J. (1999). Acoustic evidence for dynamic formant trajectories in Australian English vowels. *J. Acoust. Soc. Am.*, 106(1):458–468.

Watson, P. J. and Munson, B. (2007). A comparison of vowel acoustics between older and younger adults. In *ICPhS XVI*, pages 561–564, Saarbrücken.

Williams, D. and Escudero, P. (2014). A cross-dialectal acoustic comparison of vowels in Northern and Southern British English. *J. Acoust. Soc. Am.*, 136(5):2751–2761.

Williams, D., Van Leussen, J.-W., and Escudero, P. (2015). Beyond North American English: modelling vowel inherent spectral change in British English and Dutch. In *18th ICPhS*, Glasgow.

World Health Organization (2012a). Ageing.

World Health Organization (2012b). Definition of an older or elderly person.

Xue, S. A. and Hao, G. J. (2003). Changes in the Human vocal tact due to aging and the acoustic correlates of speech production: a pilot study. *J Speech Lang Hear Res*, 46(3):689–701.

Yuan, J. (2013). The Spectral Dynamics of Vowels in Mandarin Chinese. In *INTERSPEECH*, pages 1193–1197, Lyon, France.

Yue, M., Chen, L., Zhang, J., and Liu, H. (2014). Speaker age recognition based on isolated words by using SVM. In *CCIS2014*, pages 282–286.

Zahorian, S. A. and Jagharghi, A. J. (1993). Spectral-shape features versus formants as acoustic correlates for vowels. *J. Acoust. Soc. Am*, 94(4):1966–1982.