Universidade de Aveiro
Universidade do Porto
2022

Luciana Patrícia Martins Nunes Pereira Albuquerque

Envelhecimento Vocal: Estudo acústico-articulatório das alterações de fala com a idade

Vocal Aging: Acoustic and articulatory study of speech changes with age

Universidade de Aveiro
Universidade do Porto
2022

Luciana Patrícia Martins Nunes Pereira Albuquerque

Envelhecimento Vocal: Estudo acústico-articulatório das alterações de fala com a idade

Vocal Aging: Acoustic and articulatory study of speech changes with age

Tese apresentada à Universidade de Aveiro para cumprimento dos requisitos necessários à obtenção do grau de Doutor em Gerontologia e Geriatria, realizada sob a orientação científica da Doutora Catarina Alexandra Monteiro de Oliveira, Professora adjunta da Escola Superior de Saúde da Universidade de Aveiro, e co-orientação do Doutor António Joaquim da Silva Teixeira, Professor associado com agregação do Departamento de Eletrónica, Telecomunicações e Informática da Universidade de Aveiro e da Doutora Daniela Maria Pias de Figueiredo, Professora Coordenadora da Escola Superior de Saúde da Universidade de Aveiro.

Financial support by the Portuguese Foundation for Science and Technology (FCT) with doctoral grants SFRH/BD/115381/2016 and COVID/BD/151744/2021 are acknowledged. These are funded by Programa Operacional do Capital Humano (POCH), CENTRO2020, and by European Social Fund (FSE) under the III community support framework and national funds of the MCTES.

To my husband and my daughter, Daniel and Leonor.

o júri / the jury

presidente / president	Prof. Doutor Joaquim Arnaldo Carvalho Martins Professor Catedrático da Universidade de Aveiro
vogais / examiners committee	Prof.ª Doutora Helena Gorete Silva Moniz Professora Auxiliar da Universidade de Lisboa
	Prof.ª Doutora Susana Filipa Viegas Rodrigues Professora Adjunta da Universidade do Algarve
	Prof.ª Doutora Maria Constança Leite de Freitas Paúl Reis Torgal Professora Catedrática do Instituto de Ciências Biomédicas Abel Salazar - Universidade do Porto
	Prof. Doutor Óscar Manuel Soares Ribeiro Professor Auxiliar da Universidade de Aveiro
	Prof.ª Doutora Catarina Alexandra Monteiro de Oliveira Professora Adjunta da Universidade de Aveiro (orientadora)

agradecimentos / acknowledgements

I would like to express my immense gratitude to everyone who supported and guided me during this incredible journey. This research project has covered almost five years and was made possible thought the help and encouragement from a number of people, to whom I would like to express special thanks.

Em primeiro lugar, gostaria de agradecer à minha orientadora, Professora Catarina Oliveira, a quem devo o mais importante - o gosto que despertou em mim pela investigação, nomeadamente na área da produção de fala, e o incentivo que me deu para enveredar por este ciclo de estudos. Agradeço, ainda, a sua cuidadosa orientação científica, dedicação e acompanhamento ao longo deste percurso. Obrigada pelas constantes palavras de amizade, incentivo e coragem.

Ao Professor António Teixeira, coorientador, agradeço, não só o seu trabalho de orientação científica, mas tudo o que me ensinou e proporcionou ao longo deste percurso (sobretudo os desafios que me lançou - desde "aprender" MATLAB, Java, R, entre outros - e nos quais espero ter tido aproveitamento!). Obrigada pelas constantes palavras de incentivo, e por toda a paciência e compreensão demonstradas nos momentos mais difíceis.

À minha coorientadora, Professora Daniela Figueiredo, desejo manifestar um especial agradecimento pela sua cuidadosa orientação e pela forma pronta com que sempre respondeu aos meus pedidos.

Ao Professor Plínio Barbosa agradeço a colaboração na extração dos dados suprassegmentais. Ao Professor Pedro Sá Couto agradeço toda a disponibilidade e apoio na análise estatística.

Apresento um agradecimento especial aos colegas do Laboratório 1.10: Ana Martins, Ana Rocha, Ana Rita Valente, Bernardo Marques, David Ferreira, Diogo Cunha, Fábio Barros, Maksym Ketsmur, Mário Rodrigues, Nuno Almeida e Tiago Henriques. Obrigada pelo excelente ambiente de trabalho, que é meio caminho andado para o sucesso, e por todos os momentos de diversão que partilhamos.

Dirijo, ainda, um agradecimento a todos os Professores e colegas do projeto VoxScenes, que de alguma forma também me ajudaram a concluir este ciclo. Agradeço, especialmente, à Ana Rita Valente, ao Fábio Barros e ao Samuel Silva por todas as horas passadas na sala de ultrassons e no Zoom a reunir e a resolver problemas! Obrigada pelo vosso incansável apoio e amizade.

A todos as pessoas que colaboraram como informantes, obrigada pela disponibilidade e altruísmo. Deixo, ainda, um agradecimento às instituições que me acolheram na recolha de dados: Universidade sénior de Oliveira do Bairro, Universidade Sénior de Cacia, Academia de Saberes de Aveiro, Universidade Sénior do Centro Social Paroquial da Nossa Senhora da Nazaré, Santa Casa da Misericórdia de Oliveira do Bairro e Centro Social de Oiã.

Por fim, agradeço a toda a minha família, pelo apoio, carinho, confiança e força que me transmitiram durante todo o processo de realização deste trabalho. Em especial ao meu marido Daniel, pela sua ajuda, compreensão e incentivo constante, por acreditar em mim e aceitar as minhas ausências. À minha filha Leonor, que mesmo sem saber, me dá força para continuar.

palavras-chave

envelhecimento; vogais; suprassegmental; fala; Português Europeu; acústica de fala; produção de fala; imagens de ultrassom 2D da língua

resumo

Contextualização: Embora o processo de envelhecimento cause alterações específicas no sistema de produção de fala, o conhecimento sobre os efeitos da idade na fala é ainda disperso e incompleto. **Objetivo:** Proporcionar uma visão mais ampla das alterações segmentais e suprassegmentais da fala relacionadas com a idade no Português Europeu (PE), considerando outros aspetos, para além das características acústicas estáticas, tais como dados dinâmicos e articulatórios. **Método:** Foram criadas duas bases de dados, com dados de fala de adultos nativos do PE, obtidos através de procedimentos padronizados de gravação e segmentação: i) uma base de dados acústica contendo todas as vogais orais do PE em contexto semelhante (leitura de palavras), e também uma amostra de fala semiespontânea (descrição de imagem) produzidas por uma larga amostra de indivíduos entre os 35 e os 97 anos; ii) e outra com dados articulatórios (imagens de ultrassom da língua sincronizadas com o sinal acústico) de todas as vogais orais do PE produzidas em contextos semelhantes (pseudopalavras e palavras isoladas) por adultos de duas faixas etárias ([21-35] e [55-73]). **Resultados:** Tendo em conta as bases de dados curadas, foi analisado o efeito da idade em diversas características da fala. Acusticamente, a fala de pessoas mais velhas é caracterizada por: 1) vogais mais longas (ambos os sexos); 2) tendência para F0 diminuir nas mulheres e aumentar ligeiramente nos homens; 3) diminuição da frequência dos formantes das vogais nas mulheres; 4) redução significativa do espaço acústico das vogais nos homens; 5) vogais com maior inclinação da trajetória de F1 (ambos os sexos); 6) descrições mais curtas e com maior tempo de pausa nos homens; 7) aumento da velocidade articulatória e da velocidade de fala nas mulheres; e 8) diminuição do HNR na fala semiespontânea em mulheres. Além disso, os idosos tendem a apresentar mais sintomas depressivos que podem afetar a quantidade de fala produzida. Em relação aos dados articulatórios, a língua tende a apresentar-se mais alta e avançada em quase todas as vogais com a idade, ou seja o espaço articulatório das vogais tende a ser maior, mais alto e avançado nas mulheres mais velhas. **Conclusão:** Este estudo fornece novos dados sobre o efeito da idade na fala para uma língua diferente do inglês. Os resultados corroboram que a fala sofre alterações com a idade, que diferem em função do género, sugerindo ainda que os falantes podem desenvolver ajustes articulatórios específicos com a idade.

keywords

aging; vowels; suprasegmental speech; European Portuguese; speech acoustic; speech production; 2D-ultrasound tongue imaging

abstract

Background: Although the aging process causes specific alterations in the speech organs, the knowledge about the age effects in speech production is still disperse and incomplete. **Objective:** To provide a broader view of the age-related segmental and suprasegmental speech changes in European Portuguese (EP), considering new aspects besides static acoustic features, such as dynamic and articulatory data. **Method:** Two databases, with speech data of Portuguese adult native speakers obtained through standardized recording and segmentation procedures, were devised: i) an acoustic database containing all EP oral vowels produced in similar context (reading speech), and also a sample of semi-spontaneous speech (image description) collected from a large sample of adults between the ages 35 and 97; ii) and another with articulatory data (ultrasound (US) tongue images synchronized with speech) for all EP oral vowels produced in similar contexts (pseudowords and isolated) collected from young ([21-35]) and older ([55-73]) adults. **Results:** Based on the curated databases, various aspects of the aging speech were analyzed. Acoustically, the aging speech is characterized by: 1) longer vowels (in both genders); 2) a tendency for F0 to decrease in women and slightly increase in men; 3) lower vowel formant frequencies in females; 4) a significant reduction of the vowel acoustic space in men; 5) vowels with higher trajectory slope of F1 (in both genders); 6) shorter descriptions with higher pause time for males; 7) faster speech and articulation rate for females; and 8) lower HNR for females in semi-spontaneous speech. In addition, the total speech duration decrease is associated to non-severe depression symptoms and age. Older adults tended to present more depressive symptoms that could impact the amount of speech produced. Concerning the articulatory data, the tongue tends to be higher and more advanced with aging for almost all vowels, meaning that the vowel articulatory space tends to be higher, advanced, and bigger in older females. **Conclusion:** This study provides new information on aging speech for a language other than English. These results corroborate that speech changes with age and present different patterns between genders, and also suggest that speakers might develop specific articulatory adjustments with aging.

CONTENTS

LIST OF FIGURES

List of Tables

This page intentionally contains only this sentence.

ACRONYMS

AAA Articulate Assistant Advanced software.

ASR Automatic Speech Recognition.

C0 First DCT coefficient: formant trajectory mean.

C1 Second DCT coefficient: formant trajectory slope.

C2 Third DCT coefficient: curvature of the formant trajectory.

DCT Discrete Cosine Transform.

DEI Dorsum Excursion Index.

EBT Ensemble Bagged Trees.

EMA Electromagnetic Articulography.

EP European Portuguese.

EPG Electropalatography.

F0 Fundamental Frequency.

F1 First Formant Frequency.

F1RR First Formant Range Ratio.

F2 Second Formant Frequency.

F2RR Second Formant Range Ratio.

F3 Third Formant Frequency.

F4 Four Formant Frequency.

FCR Formant Centralization Ratio.

fgSVM Fine Gaussian Support Vector Machines.

FT Fine Tree.

HADS Hospital Anxiety and Depression Scale.

HADS-A Anxiety subscale of the Hospital Anxiety and Depression Scale.

HADS-D Depression subscale of the Hospital Anxiety and Depression Scale.

HNR Harmonics-to-Noise Ratio.

IPA International Phonetic Alphabet.

KNB Kernel Naïve Bayes.

LAOC Length of the Anterior Oral Cavity.

LPC Linear Predictive Coding.

LPTS Length of the Posterior Tongue Surface.

MFCCs Mel-Frequency Cepstrum Coefficients.

NHR Noise-to-Harmonic Ratio.

PETG Polyethylene Terephthalate Glycol-modified.

QDA Quadratic Discriminant Analysis.

RTMRI Real-Time Magnetic Resonance Imaging.

Speaking F0 Speaking Fundamental Frequency.

Spectral roc Spectral rate of change.

TA Tongue Advancement.

TCPI Tongue Constraint Position Index.

TH Tongue Height.

TL Trajectory Length.

US Ultrasound.

VAI Vowel Articulation Index.

VSA Vowel Space Area.

PART I

INTRODUCTION

This page intentionally contains only this sentence.

CHAPTER 1

THESIS OVERVIEW

1.1 Motivation

The older population is quickly increasing worldwide. Demographic aging, while due primarily to lower fertility, also reflects a human success story of increased longevity (He, Goodkind, & Kowal, 2016). However, increasing longevity led to new challenges, such as a pressure on health care costs, achieving life expectancy in good health, and living independently (He et al., 2016; Makiyama & Hirano, 2017). Portugal is one of the developed countries with the highest rate of older population (between 1970 and 2021, the percentage of people aged 65 and over increased from 9.7% to 23.4%) (Statistics Portugal, 2015, 2021), and it is classified as one of the fastest aging countries by the Organization for Economic Co-operation and Development (OECD) (Rudnicka et al., 2020). Moreover, the old-age dependency rate may increase between 2018 and 2080, from 33.9 to 67.8 older people per 100 potentially active people (Statistics Portugal, 2019; United Nations, Department of Economic and Social Affairs, Population Division, 2019).

The natural and inexorable process of aging involves changes at different levels (e.g., physiological, cognitive, psychological, social) (Makiyama & Hirano, 2017) and the human speech production mechanism is no exception (Linville, 2001; Mautner, 2011; Schötz, 2006). The anatomical and physiological changes in the speech organs (e.g. decreased lung capacity; ossification and calcification of the laryngeal cartilages; vocal fold atrophy and motor function decrease of resonant organs)(Linville & Rens, 2001; Makiyama & Hirano, 2017; Schötz, 2006), are reflected in the variation of several acoustic parameters (Eichhorn, Kent, Austin, & Vorperian, 2018; Linville & Rens, 2001; Schötz, 2006). Nonetheless, the magnitude of the speech changes depends upon the individual, as the voice is intricately linked to the dynamics of the speech organs (Makiyama & Hirano, 2017).

Unlike other languages, in which speech variations related to age have been widely studied since the 1960s (Schötz, 2006), for the European Portuguese (EP) there are only a few studies about segmental and suprasegmental changes motivated by aging (Albuquerque et al., 2014; Guimarães & Abberton, 2005; Pellegrini et al., 2013; Rato, Rodrigues, & Varanda, 2017). Additionally, there is a paucity of literature on EP oral vowels production, as the available data were collected mainly through acoustic studies (Albuquerque et al., 2014; F. Costa, 2004; Escudero, Boersma, Rauber, & Bion, 2009; M. R. D. Martins, 1973; Oliveira, Cunha, et al., 2012; Pellegrini et al., 2013; Rato et al., 2017) or in articulatory studies of nasal vowels (Cunha et al., 2019; Oliveira, Martins, Silva, & Teixeira, 2012; Oliveira, Martins, & Teixeira, 2009). In addition, there are also a few studies on the effects of aging on tongue position during speech production for other languages (Belmont, 2015; De Decker & Mackenzie, 2017; Hermes, Mertens, & Mücke, 2018; Neel & Palmer, 2012).

As several parameters of the speech signal undergo changes during the natural process of aging (Xue & Hao, 2003), and the acoustic models needed to recognize speech are typically trained using speech collected from younger adults, speech technologies have more difficulties dealing with older speech

(Baba, Yoshizawa, Yamada, Lee, & Shikano, 2004; Hämäläinen et al., 2014, 2012; Pellegrini et al., 2013; Pellegrini, Hedayati, Trancoso, Hämäläinen, & Dias, 2014). Because speech is the most natural and easy way to interact with computers (A. Teixeira et al., 2009), systems with speech based interactions can be particularly useful for older people with mobility restrictions and visual impairment (A. Teixeira et al., 2009; V. Teixeira et al., 2012; Vipperla, Renals, & Frankel, 2010). Furthermore, as most older people would like to live in their own homes as long as possible, there is an increased interest in technologies adapted to their needs to ensure their autonomy (Pellegrini et al., 2013; A. Teixeira et al., 2009; V. Teixeira et al., 2012). However, due to the decrease in performance of the automatic speech recognition (ASR) systems, older adults have more difficulties in maintaining the interaction with computers. This represents a barrier to their access to new technologies (Hämäläinen et al., 2012; V. Teixeira et al., 2012), and puts these people at risk of social isolation and exclusion. Thus, directly or indirectly, knowledge of how speech changes with age is essential for the development of ASR systems suitable for older voices (e.g. personalized reading aids and voice prostheses) (Vipperla et al., 2010).

Additionally, speech changes with age may reduce the ease of communication and affect the quality of life by limiting social interaction (Mautner, 2011) and affecting the mood (Roy, Stemple, Merrill, & Thomas, 2007; Sataloff, Kost, & Linville, 2017).

A deeper knowledge of how speech changes with age is also crucial for clinical assessment and treatment of different speech disorders, which are often age-related (Caruso, Mueller, & Shadden, 1995; Johns III, Arviso, & Ramadan, 2011), and to provide information for other fields of knowledge (e.g., phonetics, speech science, forensic linguistics and biometric recognition) (Lanitis, 2010; Morrison, 2013).

1.2 Thesis statement

Speech changes with age are complex, as they are largely dependent on the speaker, as well as there are substantial gender differences in the extent and timing of this process. Taking this into account, this thesis should provide a clear picture of the age-related segmental and suprasegmental speech changes in EP, allowing to push the state-of-the-art of the lifespan speech changes and bearing in mind new aspects besides static acoustic features, such as dynamic and articulatory data. In the acoustic point of view, it is important to go beyond the traditional static analysis, using dynamic data from the vowel formants. Thus, a dynamic approach to study the aging effect on vowel production could be important, since a static approach only reflects anatomical differences among speakers. Concerning the articulatory aspects, it is essential to understand the relationship between articulatory gestures and acoustic changes with age.

Additionally, this study adds to the growing body of data on the effects of age on the acoustic (and articulatory) properties of speech, providing information of different levels from adults who speak a language different from English. In that sense, it might help to better understand cross-linguistic similarities and language-particular features of aging.

1.3 Objectives

Motivated by the need to increase knowledge about the effect of age on speech production, this thesis focuses on healthy adults [1] and the main goal of this work is to perform a comprehensive analysis of age and gender effects in EP speech at segmental and suprasegmental level using acoustic and articulatory data.

The following objectives are then defined in order to achieve this main goal:

- to analyze the age effects on duration, fundamental frequency (F0) and formant frequencies (F1, F2 and F3) for all EP oral vowels produced by a group of speakers of both genders;

- to investigate the age-related effect on dynamic characteristics of vowel formants;

- to study the effects of aging on suprasegmental acoustic parameters (e.g., speech rhythm, speaking fundamental frequency (speaking F0) and Harmonics-to-Noise Ratio (HNR)) using semi-spontaneous speech;

- to explore the association between the acoustic parameters (at segmental and suprasegmental level) and depressive and anxiety symptoms considering age;

- to evaluate the age-related changes in speech articulation, namely in the tongue contours, in EP oral vowels produced by healthy speakers using ultrasound (US) imaging.

1.4 Methodology

Considering the lack of speech data on EP, the development of speech databases, with acoustic and articulatory data of healthy adults using standardized recording procedures, is essential. The collected data will be used on acoustic and articulatory studies of speech over the lifespan.

In general, the overall methodology employed in the development of the databases and the subsequent studies is based on cross-sectional data. Thus, a cross-sectional study is a type of observational study, which is limited to a single evaluation moment (Fortin & Salgueiro, 1999). Therefore, there is no follow-up period for the individuals in the sample. The participants' selection follows certain criteria and it is a sampling for convenience (i.e., it is a non-casual sampling) (Fortin & Salgueiro, 1999). In this observational cross-sectional study, the acoustic and articulatory parameters are measured from the speech data and the independent variables are obtained through self-reported questionnaires, which are structured measuring instruments (Fortin & Salgueiro, 1999).

The detailed description of the research design (i.e, speakers selection, speech corpus, materials, data collection, data analysis) is presented in Part II: Methods. Each specific study also specifies the rationale for choosing the method.

[1] Note that, in this document "healthy adults" are understood to be individuals with no history of impairments that affect the voice, speech or language.

1.5 Contributions

This thesis was funded by the Fundação para a Ciência e Tecnologia (FCT) from September 1, 2017 to August 31, 2022 [Grant numbers: SFRH/BD/115381/2016 and COVID/BD/151744/2021] and was conducted in the Institute of Electronics and Informatics Engineering of Aveiro (IEETA) and Center for Health Technology and Services Research (CINTESIS).

The present work allowed the development of two speech databases for EP: i) one acoustic database with 144 speakers who not only produced disyllabic words in a carrier sentence, with the EP vowels [i], [e], [ɛ], [a], [o], [ɔ] and [u] in stressed position and the vowels [ɨ] and [ɐ] in unstressed position, but also a semi-spontaneous speech sample; ii) and another one with articulatory data for all EP oral vowels with US, produced by 31 adult speakers of two different age groups ([21-35] and [55-73]). The databases were developed in partnership with the VoxSenes project, and they could also be explored in different ways in future studies.

In addition, the outcomes in the context of this thesis, or related with it, can be grouped into two main areas: 1) age related changes on EP speech acoustics; 2) articulatory analysis of the age effect on EP vowel production. The dissemination of these achievements was made through the publication and presentation of scientific papers in peer-reviewed conferences and journals, as described below by main area.

1.5.1 Age-related changes on EP speech acoustics

The preliminary results of the acoustic speech database analysis regarding the effects of age on vowel acoustics were presented in:

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

2. **Albuquerque, L.**, Oliveira, C., Teixeira, A., Sa-Couto, P., & Figueiredo, D. (2019). Efeito da idade nas vogais orais do Português Europeu. In *XXXV Encontro Nacional da Associação Portuguesa de Linguística*. Braga, Portugal: APL. (oral presentation)

   The deep analysis of the age-related effect on vowel acoustics using static measures was presented in:

3. **Albuquerque, L.**, Oliveira, C., Teixeira, A., Sa-Couto, P., & Figueiredo, D. (2020). A comprehensive analysis of age and gender effects in European Portuguese oral vowels. *Journal of Voice*, (*in press*).

   The usefulness of dynamic features for classification tasks (i.e., age classification) and characterization of vowel acoustics was presented in:

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

5. **Albuquerque, L.**, Oliveira, C., Teixeira, A., & Figueiredo, D. (2021). Eppur si muove: Formant dynamics is relevant for the study of speech aging effects. In *14th BIOSTEC* (pp. 276–283). Online.

6. **Albuquerque, L.**, Teixeira, A., Oliveira, C., & Figueiredo, D. (2022). Age and vowel classification improvement by the inclusion of vowel dynamic features. *International Journal of Speech Technology*, 25(1025–1040).

   The effects of age on suprasegmental acoustic parameters were presented in:

7. **Albuquerque, L.**, Valente, A. R. S., Teixeira, A., Oliveira, C., & Figueiredo, D. (2021). Acoustic changes in spontaneous speech with age. In *VIII Congreso Internacional de Fonética Experimental*. Girona, Spain. (oral presentation)

8. **Albuquerque, L.**, Valente, A. R. S., Teixeira, A., Oliveira, C., & Figueiredo, D. (2021). Age and gender effects in European Portuguese spontaneous speech. *Loquens*, 8(1–2), e077.

   The relationship between acoustic speech features (i.e., segmental and suprasegmental level) and non-severe levels of anxiety and depression symptoms considering age was presented in:

9. **Albuquerque, L.**, Valente, A. R. S., Teixeira, A., Figueiredo, D., Sa-Couto, P., & Oliveira, C. (2021). Association between acoustic speech features and non-severe levels of anxiety and depression symptoms across lifespan. *PLOS ONE*, 16(4), e0248842.

1.5.2 Articulatory analysis of the age effect on EP vowel production

In order to simplify the time-consuming task of US image segmentation, a method of extraction and analyzing the image of the tongue contours was developed with our collaboration:

1. Barros, F., Valente, A. R., **Albuquerque, L.**, Silva, S., Teixeira, A., & Oliveira, C. (2020). Contributions to a quantitative unsupervised processing and analysis of tongue in ultrasound images. *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 12132 LNCS, 170–181.

2. Barros, F., Silva, S., **Albuquerque, L.**, Valente, A. R., Teixeira, A., Martins, P., & Oliveira, C. (2020). Towards the use of ultrasonography to study aging effects in vowel production. In *12th ISSP*, Online.

   Before processing all the speech articulatory database, a pilot study only with data of four females, two younger and two older, was presented in:

3. **Albuquerque, L.**, Valente, A. R., Barros, F., Teixeira, A., Silva, S., Martins, P., & Oliveira, C. (2021). The age effects on EP vowel production: an ultrasound pilot study. In *IberSPEECH 2021* (pp. 245–249). Valladolid, Spain: ISCA. (oral presentation)

   The full analysis of the age-related articulatory differences in EP vowels using US imaging was presented in:

4. **Albuquerque, L.**, Valente, A. R., Barros, F., Teixeira, A., Silva, S., Martins, P., & Oliveira, C. (2022). Exploring the age effects on European Portuguese vowel production: An ultrasound study. *Applied Sciences*,12(3),1396.

1.5.3 Other contributions

Notwithstanding the exclusivity agreement associated with the grant SFRH/BD/115381/2016 awarded by the FCT, the funder have formally authorized the participation of the awardee in other funded research project - VoxSenes Project - in parallel with the execution of the PhD work plan. The following publications resulted from this collaboration. Although these scientific outcomes are not part of the thesis structure, those were of relevance to improve the candidate's knowledge on the field.

1. Oliveira, C., Valente, A. R., **Albuquerque, L.**, Barros, F., Martins, P., Silva, S., & Teixeira, A. (2021). The VoxSenes project: A study of segmental changes and rhythm variations on European Portuguese aging voice. In *IberSPEECH 2021* (pp. 135-138). Valladolid, Spain: ISCA.

2. Oliveira, C., **Albuquerque, L.**, & Valente, A. R. (2021). A fala no idoso (Speech in older adults). In Verbetes LBASS. Available in: `http://www.letras.ufmg.br/lbass/`

3. Valente, A. R., Oliveira, C., **Albuquerque, L.**, Teixeira, A., & Barbosa, P. A. (2021). Prosodic Changes with Age: A Longitudinal Study on a Famous European Portuguese Native Speaker. *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 12997 LNAI. 726–736.

4. Valente, A. R., Oliveira, C., **Albuquerque, L.**, Teixeira, A., & Barbosa, P. A. (submitted). Prosodic changes with age: A Longitudinal Study with public figures in European Portuguese. *Speech Communication*.

1.6 Structure

In this thesis, there are four main parts, listed and described in the following lines:

I. Introduction: this part includes three chapters. The present chapter (Chapter 1) contains the thesis overview and the motivation and objectives of the present work that led to the development of the speech databases and the speech analysis. Chapter 2 describes the background related with aging and it reviews the anatomical and physiological changes behind the speech alterations with age. Finally, Chapter 3 presents an overview of the acoustic and articulatory changes with age, as well as it describes the impact of anxiety and depression symptoms on speech acoustics.

II. Methods: this part includes three chapters. Chapter 4 describes the method used to obtain the acoustic speech database of EP. Chapter 5 reports the extraction of the acoustic measures at both segmental and suprasegmental levels, and also presents the sociodemographic characteristics of the speakers. Chapter 6 reports the method used to devise the articulatory database of EP oral vowels.

III. Acoustic and Articulatory Studies: here are included the main results obtained with the analysis of the databases devised. The main body of this part is structured into five chapters, reporting the studies (published or accepted to international indexed journals) that have been conducted throughout the project to accomplish the defined objectives. Chapter 7 contains the age-related acoustic changes in EP oral vowels. Chapter 8 presents a study about age and vowel classification improvement by the inclusion of vowel dynamic features. Chapter 9 includes the age-related acoustic changes in semi-spontaneous speech. The association between speech acoustics parameters and mood symptoms with age is presented in Chapter 10. Chapter 11 describes the articulatory results about the age effect on EP oral vowel production. All the chapters of this part contain: an introduction (to the problem); a brief description of the methodology; results; and a discussion.

IV. Discussion and Conclusion: finally, this part includes Chapter 12 that contains the summary of the main findings of each study, followed by an overall discussion where the results of all studies are analyzed as a whole. The study limitations and the future work conclude the document.

This page intentionally contains only this sentence.

CHAPTER 2

BACKGROUND

This chapter provides an overview of the background related to aging and reviews the anatomical and physiological changes behind the speech alterations with age.

2.1 Human aging: basic concepts and definitions

Aging is a concept that has been used in a variety of ways. The lack of a universal definition of aging arises the fact that it can be considered according to social, behavioural, physiological, morphological, cellular and molecular changes (Balcombe & Sinclair, 2001).

Biological aging, also known as senescence, is one of the most complex biological processes (J. P. Costa et al., 2016), and results from a series of lifespan-associated declines in molecular, cellular, tissue, and organ function (Broskey et al., 2019). Many theories attempt to explain the aging process, but neither of them appears to be fully satisfactory (K. Jin, 2010). Also, it is almost impossible to give a complete overview of all theories (J. P. Costa et al., 2016), but two main perspectives are advocated: program-related aging and damage-/error-related aging (Makiyama & Hirano, 2017).

The programmed or adaptive theories posit that the maximum lifespan is predetermined by the genes in each species (Hoffnung et al., 2015) (i.e., aging is the result of a pre-written program, which controls growth, development and aging (Balcombe & Sinclair, 2001)). These theories also suggest that there is an age-related deliberate deterioration since a limited lifespan results in evolutionary benefits (J. P. Costa et al., 2016; Goldsmith, 2012). Therefore, aging may represent a way to avoid overpopulation through the elimination of post-reproductive age individuals, who would compete for resources with their younger generations This also enables ongoing evolutionary adaptation by promoting succession of generations (Kirkwood & Melov, 2011). On the other hand, under identical conditions, significant lifespan differences for many species are observed, which seems to indicate that aging is not predetermined or programmed, but rather the end-result of a “wear-and-tear” (J. P. Costa et al., 2016). In this line of thought, the damage or error theories arise. Hence, aging happens because of the lack of natural selection at the post-reproductive stage of life (J. P. Costa et al., 2016). These theories highlight environmental aggressions to living organisms that lead to cumulative damage at various levels as the cause of aging (K. Jin, 2010). Nonetheless, these theories are not mutually exclusive (Troen, 2003) and interact with each other in a complex way (K. Jin, 2010). For that, more recently, a third category - combined theories - has emerged (J. P. Costa et al., 2016). In these theories the aging process is considered at a more comprehensive and global degree (J. P. Costa et al., 2016).

Overall, aging is a continuous process characterized by a progressive decline in physiological homeostasis, leading to impaired function and increased susceptibility to illness and death (Balcombe & Sinclair, 2001; Broskey et al., 2019; López-Otín, Blasco, Partridge, Serrano, & Kroemer, 2013; Troen, 2003). However, the deterioration of physiological homeostasis with the inevitable advancement of age is variable (Broskey et al., 2019), since aging is a heterogeneous event, undoubtedly very individual

(Balcombe & Sinclair, 2001; Kafková, 2016). Despite many aging theories, the underlying causes of these physiological deteriorations remain controversial (Balcombe & Sinclair, 2001; Broskey et al., 2019; Lemoine, 2020), and can be partitioned into primary or secondary aging (Broskey et al., 2019).

Primary aging refers to the deterioration in structure and function that is due to the aging process (i.e., independent of disease and environment) (Holloszy, 2000). Secondary aging refers to additional deleterious structural and functional age-related changes caused by diseases, environmental or lifestyle factors (Holloszy, 2000). In other words, primary aging describes the effect of the inevitable, yet variable, progression of aging, whereas secondary aging describes the effect of extrinsic factors such as diet, smoking and physical activity that can accelerate or decelerate the aging process (Broskey et al., 2019). In this sense, lifespan is dictated by the process of aging, and aging may occur as a result of genetic (intrinsic) and, to a greater extent, environmental (extrinsic) factors and time (Balcombe & Sinclair, 2001; Dodig, Čepelak, & Pavić, 2019). Intrinsic aging must involve at least one of these processes: progressive structural damage, functional decline or depletion, changes associated with an aging phenotype or an increasing probability of death or disease (Lemoine, 2020). Also, Lemoine (2020) described aging as a binary process, with a progressively shifting balance between mechanisms that limit the lifespan ("promotive") and mechanisms that modulate their effects ("protective"), i.e., degradation and compensation mechanisms of aging.

Furthermore, aging could be classified as psychological, social, functional or chronological aging (Chalise, 2019). Regarding the psychological aspects of aging, this process involves changes in memory, learning, intelligence, personality, and coping (Chalise, 2019). Moreover, with the increasing age rises the expectation of different roles for individuals in society (Balcombe & Sinclair, 2001). Thus, social aging refers to changes in roles and relationships as age increases (Chalise, 2019). The person's social age within a specific context can be very important because it determines the meaning of aging for the person, and this can make aging a positive or negative experience (Chalise, 2019) By the end, functional aging is related to how people compare psychologically to others of a similar age (Chalise, 2019).

Chronological aging is the number of years a person has lived so far (Chalise, 2019). In this sense aging starts at conception and continues until death. It is easy to measure and it is justified by the fact that, with the passage of time, there is an increasing incidence of physical, mental and functional problems and an increase in age-specific mortality rates (Balcombe & Sinclair, 2001). However, chronological age may not match a person's biological, psychological, or social age (Chalise, 2019), but most studies use chronological age as a proxy.

Thereby, more important than aging is healthy aging, which is the process of developing and keeping the functional ability that allows well-being in older age (World Health Organization, 2020, p.3). The functional ability is determined by the intrinsic capacity of an individual (i.e., the physical and mental capacities), the environment in which he or she lives (e.g., physical, social and policy environments) and the interactions among them (World Health Organization, 2020).

2.1.1 Life cycle

The division of the life cycle into periods is a social construction (i.e., it is a way of conceptualizing human lives from birth to death) (Papalia, Feldman, Monteiro, & Silva, 2013; Settersten & Mayer, 1997). In this sense, there is no general agreement on the age at which a person becomes old (Birren & Schaie, 2001; Papalia et al., 2013), and the World Health Organization (2017b) refers that as people age, they move along a continuum that stretches from robustness to frailty.

According to Chalise (2019), biological aging includes three stages: growth; maturation; and senescence [1] (after age 30, when the physical body begins to wear out and their functioning declines) (Chalise, 2019; Makiyama & Hirano, 2017). According to Balcombe and Sinclair (2001), there are certain landmarks that help to divide the aging process in various stages, namely birth, puberty and the menopause in women (menopause occurs around the age of 51, with 3 to 4 years interval deviation (Lã & Ardura, 2020)). Despite that the first four-five decades of life are well defined, no clear landmarks exist for either men or women after this point on, so the process of biological aging lacks clear definition (Balcombe & Sinclair, 2001). Nonetheless, researchers agree with the division of adulthood in age groups due to the large span of years, but the age division tend to be very variable.

From a developmental perspective, according to the classification of the American Psychological Association (2020), adulthood period begins after adolescence and is sometimes divided into young adulthood (around 20 to 35 years of age); middle adulthood (36 to 64); and later adulthood (65 or older). Other authors agree with the division in three phases, but suggest different age cohorts, such as: 20 – 40, 40 – 65 and 65 years old and beyond (Papalia et al., 2013); or 25 – 44, 45 – 59 and 60 or older (Dyussenbayev, 2017; Yue, Chen, Zhang, & Liu, 2014), as advocated by the World Health Organization. Although there is no set age marking that a person is old, the developed world often adopts the retirement age of 60 or 65 as the entry point for old age (Robinson, MacDonald, & Broadbent, 2014; United Nations, Department of Economic and Social Affairs, Population Division, 2019), which provides a simple, clear and easily replicable way to measure and track various indicators of population aging (United Nations, Department of Economic and Social Affairs, Population Division, 2019). Also, these ages often represent the lower age boundary that differentiated “younger” from “older” participants in the majority of studies on aging speech (Mautner, 2011). On speech communication studies, Hazan (2017) also suggests the following age division of the adulthood period: young adult (19-35 years), middle aged (36-64), and older adult (65 or older).

Due to the increasing heterogeneity of the older population, the gerontologists make the distinction among the young-old (65 to 74), the old-old or the middle old (75 to 84), and the oldest old (age 85 and beyond) (American Psychological Association, 2020; Kafková, 2016; Nosraty, 2018). Neugarten (1974) introduced her division into young-old and old-old older people with the age boundary of 75 years in 1974, and since then, life expectancy has increased significantly (Kafková, 2016). For that, according to Kafková (2016) the age boundary of 75 years, widely used in analytic practice, seems to be unsuitable since frailty and general health decline occur more often after 80 in both men and women. Thus, Kafková (2016) suggests the boundary of 80 years to define the old-old. Also, Barnes (2011a,

[1] Although “aging” is commonly used to refer the post-maturational processes, the more correct term for this is “senescence”, while “aging” can refer to any time-related process (Balcombe & Sinclair, 2001; Troen, 2003). However, in this document, “senescence” and “aging” will be used interchangeably.

2011b) divide the late adulthood into two categories: $65-80$ and 80 years old or older. However, due to the increase of longevity, other stages have been suggested, such as longevity, late old age or long-livers (the age of 90 and beyond) (Dyussenbayev, 2017; Yue et al., 2014), and the centenarians (100+) (Willcox, Willcox, & Poon, 2010).

Despite the stages being subject to considerable variation in age ranges, young-old is the period in which the changes of the human physical capabilities that occur with the aging process become more evident (Barnes, 2011b; Dodig et al., 2019; Dyussenbayev, 2017). This period generally includes good health and social engagement, functional reserve capacity, knowledge and expertise, and adaptive flexibility in daily life (Barnes, 2011b). Then, followed by periods where the the decline of human physical abilities are more pronounced (Dodig et al., 2019; Dyussenbayev, 2017).

2.2 Aging of the speech production system

Aging, as an inevitable biological process, impacts the human body in a plethora of ways (Hermes et al., 2018). It is generally reported that the human peaks at the age of 30, and then declines about 1% per year in not only physiological aspects, but also in psychological and sociological aspects (Chalise, 2019; Makiyama & Hirano, 2017; Ramig et al., 2001). Age-related changes occur in any part of the body, including musculoskeletal, cardiovascular, respiratory, neurological, gastrointestinal, immunological, endocrine, and dermal systems (Hermes et al., 2018; Makiyama & Hirano, 2017). Some of the age-related body changes decrease active life expectancy, while others have few or no health consequences (e.g., wrinkles and sagging skin, gray hair or hair loss) (Quadagno, 2017).

Age-related changes take place in different tissues and organs, and the speech production system is no exception (Hermes et al., 2018; Makiyama & Hirano, 2017). This is a complex system, as there are no organs which function is only speech production (Barbosa & Madureira, 2015; Lieberman & Biumstein, 1988). In other words, the organs that allow the production of speech (sounds) can be seen as an adaptation of the respiratory and digestive systems (Lieberman & Biumstein, 1988).

The major anatomical components of the speech production system can be divided into three parts: the subglottal component (i.e., the lungs and associated respiratory musculature), the larynx (i.e., the vocal folds and glottis) and the supraglottal cavities (i.e., the pharynx, oral tract and nasal tract) (Lieberman & Biumstein, 1988; Linville, 2001; Redford, 2015; Titze, 1994). In general, the subglottal system supplies energy (air flow), and the laryngeal and supraglottal structures are responsible for its modulation, in order to produce speech sounds (Barbosa & Madureira, 2015; Lieberman & Biumstein, 1988; Redford, 2015). The speech production system primarily consists of the respiratory, phonatory, and articulatory systems, but also involves the whole body, mood, cognition and the neuromuscular system (Makiyama & Hirano, 2017; Mendes, Guerreiro, Simões, & Moreira, 2013). Thus, the motor function of the speech production system envelops efficiency, coordination, and precision (Titze, 1994). Moreover, speech production is determined by genetic, linguistic, social and cultural factors (Mendes et al., 2013).

From young adulthood to old age, the speech production mechanism undergoes numerous anatomical and physiological changes that are expected to play an important role in speech production (Goozée, Stephenson, Murdoch, Darnell, & Lapointe, 2005). Furthermore, the timing of changes on speech may vary across individuals (Fuchs, Koenig, & Gerstenberg, 2021), as well as there are substantial gender differences in the extent and timing of the aging process (Linville, 2001; Makiyama & Hirano, 2017; Mautner, 2011; Schötz, 2006).

There are some important periods of vocal change, such as early childhood, puberty, and advanced age (Titze, 1994). However, the primary purpose of the following sections is to review the normal anatomophysiological changes associated with aging on the respiratory, laryngeal, articulatory and neuromuscular systems for a better understanding of their effects on aging speech. Thus, this review is not meant to focus on childhood and puberty vocal changes, neither on age-related diseases.

2.2.1 Respiratory system

The "breathing for speech" is intricately related to respiration, since speech is an overlaid function (Hooper & Cralidis, 2009). Moreover, the respiratory system plays a crucial role in the speech production process, since it is the "power source" of the voice (Desjardins, Halstead, Simpson, Flume, & Bonilha, 2022; Linville, 2001; Sataloff et al., 2017). The respiratory muscles are also central to the airflow management by counteracting the recoil pressures of the lungs to control the expiratory phase of speech, which allows the production of longer utterances with constant pressure (Desjardins et al., 2022).

The subglottal respiratory system comprises the trachea, thoracic cage, lungs (mainly composed by bronchi, bronchioles, and alveoli), respiratory muscles and diaphragm (Barbosa & Madureira, 2015; Sharma & Goodwin, 2006). Anatomical age-related changes in these structures lead to changes in the physiology of breathing, which affect the speech breathing as well as the voice produced by older speakers (Linville, 2001; Redford, 2015; Schötz, 2006).

The respiratory system reaches its full potential as the chest enlarges and thoracic and abdominal musculature strengthens, which occur usually during young adulthood (Sataloff et al., 2017). The maximal function of the respiratory system is reached at approximately the age of 20 for females and 25 for males (Janssens, Pache, & Nicod, 1999; Linville, 2001), and it remains steady with very minimal change until the age 35 (Sharma & Goodwin, 2006). After, the respiratory function declines with increasing age in both men and women due to three factors: increased stiffness of the thoracic cage, decreased strength of the respiratory muscles, and decreased lung capacity (mainly due to loss of lung elasticity) (Janssens, 2005; Janssens et al., 1999; Linville, 2001; Redford, 2015; Sataloff et al., 2017; Schötz, 2006).

Normal structural changes with age in the thorax include stiffening of the thoracic cage from calcification of the costal and intercostal cartilages, and arthritis of the costovertebral and chondrosternal joints (Aalami, Fang, Song, & Nacamuli, 2003; Janssens, 2005; Sharma & Goodwin, 2006; Thompson & Chen, 2021). These alterations result in less movement in response to respiratory muscles force (Linville, 2001). Thus, as the chest wall stiffness increases with aging, compliance decreases (Aalami et al., 2003; Janssens, 2005). The shape of the thorax also changes between young adulthood and old age (Janssens, 2005; Linville, 2001).

The diaphragmatic and inter-costal muscle strength has been reported to decrease with age, mainly in men (Aalami et al., 2003; Linville, 2001; Sharma & Goodwin, 2006; Thompson & Chen, 2021).

One the other hand, the inspiratory muscle strength and the inferior expansion of the chest wall may be more impaired in women with aging (Redford, 2015). The rate of the respiratory muscles contraction also tends to decrease with advancing age (Sataloff et al., 2017). The age-related decrease in strength of the respiratory muscles may occur due to muscle atrophy and a decrease in fast twitch fibers with age, responsible for generating higher peak tensions (Sharma & Goodwin, 2006). The intercostal muscles mass starts to decrease after the age of 50, with a greater impact on the expiratory muscles (Lalley, 2013).

With aging the lungs and bronchi lower within the thorax, and the lungs tend to be smaller, lighter, and less elastic (Linville, 2001). The most significant structural age-related change in the lungs is the loss of elasticity of lung tissue (Linville, 2001) (that tends to be greater in men than in women, particularly between ages 45 and 58 (Redford, 2015)), resulting in lower static recoil pressures (Redford, 2015). The elastic fibers around the alveolar duct start to degenerate around the age 50 (Aalami et al., 2003; Sharma & Goodwin, 2006). The respiratory bronchioles and alveolar ducts increase in size significantly with aging (Aalami et al., 2003; Linville, 2001), mainly after 60 years of age (Aalami et al., 2003). It was also observed a fusion of alveoli and thickening of blood vessel walls in older adults (Linville, 2001).

All these normal aging structural changes lead to functional losses (Sataloff et al., 2017), but the decline in respiratory function during adulthood tends not to be linear and accelerates with aging (Janssens, 2005; Linville, 2001). The most important age-related functional losses include: decreased elastic recoil of lung tissues, reduced vital capacity, increased residual volume, decreased respiratory (i.e., expiratory and inspiratory) reserves, and reduced expiratory flow rates (Aalami et al., 2003; Hoit & Hixon, 1987; Hoit, Hixon, Altman, & Morgan, 1989; Janssens, 2005; Linville, 2001; Ptacek, Sander, Maloney, & Jackson, 1966; Redford, 2015; Sataloff et al., 2017; Thompson & Chen, 2021). In general, pulmonary function tests have shown greater rates of decline after the age of 65 (Knudson et al., 1983, as cited in Aalami et al., 2003), as well as higher decrease in men than in women as age increases (Aalami et al., 2003; Linville, 2001). These changes in respiratory function impact speech breathing and can also affect subglottal pressure, duration of phonation, pitch, and loudness of the voice (Makiyama & Hirano, 2017; Sataloff et al., 2017). Specifically, older adults tend to start and finish speech at higher lung volumes, use greater lung volume excursions, and use a higher percent of their lung volume per syllable. (Hoit & Hixon, 1987; Hoit et al., 1989; Huber, 2008; Huber & Spruill III, 2008; Redford, 2015; Sperry & Klich, 1992).

2.2.2 Phonatory system

The second stage of speech production is phonation and it occurs in the laryngeal area (Ball, 2021; Lieberman & Biumstein, 1988). The phonation process refers to the vibratory actions of the two vocal folds, which are soft tissue structures located in the larynx (Ball, 2021; Redford, 2015). The larynx is a structure positioned at the top of the trachea that is composed of four anatomical units: mucosa, skeleton, intrinsic muscles, and extrinsic muscles (Ball, 2021). The vocal folds are a complex structure that consists of five layers (i.e., the epithelium, the three layers of the lamina propria and the thyroarytnoid muscle), and each layer has different mechanical properties, which allow their proper vibration (Linville, 2001; Sataloff, 1991; Titze, 1994). The extrinsic laryngeal musculature maintains the position of the larynx in the neck, while the intrinsic muscles are responsible for adduction, abduction, and longitudinal tension of the vocal folds (Sataloff, 1991).

Succinctly, the flow of air arising from the lungs overcomes the resistance of the vocal folds and moves them, starting the vibratory process (Desjardins et al., 2022). Therefore, a triangular space between the vocal folds, known as the glottis, is opened and closed rapidly during speech production, generating a source of acoustic energy that characterizes the phonation process (Ball, 2021; Lieberman & Biumstein, 1988; Redford, 2015). In addition to phonation, there are other basic laryngeal functions: airway opening for respiration; and airway protection that is extremely important during the process of swallowing (Hardcastle, Laver, & Gibbon, 2012).

The most dramatic laryngeal changes occur during the transition from childhood to adolescence. However, the larynx also undergoes extensive anatomical and physiological changes after it has reached its full size in puberty (Cox & Selent, 2015; Sataloff et al., 2017; Schötz, 2006). Although some conflicting data has been reported, age-related changes in the laryngeal structures during adulthood are common and exhibit substantial gender differences, appearing to occur earlier and in a more extensive way in men than in women (Kost & Sataloff, 2020; Linville, 2001; Pontes, Brasolotto, & Behlau, 2005; Schötz, 2006), as summarized in Table 2.1. Specifically, in older females, thickening and edema of the vocal folds, protuberant vocal process, increase of the glottal contact during phonation and tremor of laryngeal structures were the most observed features (Caruso et al., 1995; Lenell, Sandage, & Johnson, 2019; Pontes et al., 2005; Pontes, Yamasaki, & Behlau, 2006; Zamponi, Mazzilli, Mazzilli, & Fantini, 2021). For males, the vocal folds tend to thin, which can contribute to bowing of the membranous portion of vocal folds, the decrease of the vocal fold contact and higher incidence of spindlle-shaped glottal gaps (Caruso et al., 1995; Lenell et al., 2019; Pontes et al., 2005, 2006; Schötz, 2006).

Generally, the age-related morphological changes in the larynx are called presbylarynx and their effect on vocal quality is called presbyphonia (i.e., age-related dysphonia) (Galluzzi & Garavello, 2018; Hagen, Lyons, & Nuss, 1996; Hooper & Cralidis, 2009; Kendall, 2007; Mallick, Garas, & McGlashan, 2019; Pontes et al., 2005; M. Santos et al., 2021; Takano et al., 2010). According to some authors (Sataloff et al., 2017; Woo, Casper, Colton, & Brewer, 1992), presbylarynx should be a diagnosis of exclusion made only after careful medical and speech evaluation. Thus, it is a diagnosis of exclusion made in absence of other laryngeal, neurological or pulmonary diseases (Arviso & Johns III, 2009; Galluzzi & Garavello, 2018; Kendall, 2007; Kost & Sataloff, 2020; Mallick et al., 2019; R. Martins et al., 2015). However, the presence of presbylarynx alone does not necessarily lead to the development of presbyphonia (Mallick et al., 2019; Pontes et al., 2005) (i.e., a gradual "weakening" of voice, hoarseness, breathiness, vocal fatigue, decrease of vocal intensity, poor vocal projection, phonatory instability, shorter phonation duration and lack of pneumophonic coordination (Arviso & Johns III, 2009; Galluzzi & Garavello, 2018; Kendall, 2007; Kost & Sataloff, 2020; Mallick et al., 2019; R. Martins et al., 2015)).

Table 2.1: Gender differences in normal structural aging of the larynx.

Laryngeal structure		Alteration	Gender differences	References
Hyoid bone		changes in hyoid bone position	more evident in males	Feng et al. (2014); Türkmen et al. (2012)
Cartilages		ossification that begins during the early decades of adulthood in the thyroid cartilage, then in the cricoid and in the arytenoids (hyaline matrix) calcification occurs later and affects epiglottis, corniculate cartilages, cuneiform cartilages and parts of the arytnoid cartilage (elastic matrix)	earlier and more extensive in males; but generally most cartilage is ossified by the age of 65	Hixon, Weismer, and Hoit (2018); Linville (2001); Mizuta (2017); Mueller (1997); Turk and Hogg (1993); Türkmen et al. (2012)
Cricoarytenoid joint		general deterioration in both their joint capsules and articular surfaces that restrict the movement of the joint	more evident in males	Hixon et al. (2018); Kost and Sataloff (2020); Linville (2001)
Intrinsic muscles		atrophy; progressive loss of muscle mass in addition to a decrease in function	gender differences is unclear	Baker, Ramig, Sapir, Luschei, and Smith (2001); Linville (2001); Mizuta (2017); Schötz (2006)
Vocal folds	Epithelium	increase of cell desquamation and more undulation; reduction in the epithelial cell density	more evident in men	Galluzzi and Garavello (2018); Gonçalves, Dos Santos, Pessin, and Martins (2016); Sataloff et al. (2017); Ximenes Filho, Tsuji, Nascimento, and Sennes (2003)
		the thickness of the epithelium progressively increase (i.e., thickening) until the age 70 in both genders	continues after in females and declines thereafter in males	Hirano, Kurita, and Sakaguchi (1989); Linville (2001); Sataloff et al. (2017); Schötz (2006)
	Superficial layer of the lamina propria	becomes thicker, more edematous, swollen with fluid, and declines in fiber density	similar in both genders	Hixon et al. (2018); Pontes et al. (2005); Sataloff et al. (2017)
		morphological changes in collagen (that migh affect the viscoelasticity of the vocal folds)	increase in males and decrease in females	Branco, Fabro, Gonçalves, and Martins (2015); Hammond, Gray, and Butler (2000); Mizuta (2017); Roberts, Morton, and Al-Ali (2011)
		hyaluronic acid level decrease	study only with males	Branco, Rodrigues, Fabro, Fonseca-Alves, and Martins (2014)
		elastic fibers appear to lose their functions due to morphologic and metabolic changes (which may limit the vibration of the tissue)	more significant in men	Branco et al. (2015); Mizuta (2017); Pontes et al. (2005); Roberts et al. (2011); I. A. Santos (1997)
	Intermediate layer of the lamina propria	thins as the elastic fibers wane in size and number (i.e., atrophy)	in males (less marked in females)	Hirano et al. (1989); Hixon et al. (2018); Ximenes Filho et al. (2003)
	Deep layer of the lamina propria	thickens as collagenous fibers increase in size and density (becoming fibrotic)	especially in males over 50 years; (not marked in females)	Branco et al. (2015); Hirano et al. (1989); Hixon et al. (2018); Pontes et al. (2005); Roberts et al. (2011)
		collagen decrease	only significant in females	Hammond et al. (2000)
		changes in the elastic fibers that lead to reduce viscoelasticity and the vocal fold pliability	no data on gender differences	Branco et al. (2015); Hammond et al. (2000); Roberts et al. (2011)
		hyaluronic acid level decrease	study only with males	Branco et al. (2014)
	Thyroarytnoid muscle	muscle atrophy/degeneration especially over 60th decade; decrease in fiber diameter; reduction in the muscle mass, endurance, speed and force of contraction	gender differences is unclear	Kersing and Jennekens (2004); Linville (2001); R. Martins et al. (2015); Thomas, Harrison, and Stemple (2008)
Macula flavae		decrease in the number and activation of fibroblasts, and less hyaluronic acid is produced, which influence the viscoelasticity and stiffness of vibrating tissue	no data on gender differences	Bruzzi et al. (2017); Linville (2001); Sato and Hirano (1996)
Conus elasticus		thinning of fibers; separation/fragmentation of fiber bundles mainly after the age of 50	in males (few changes observed in females)	Linville (2001)
Mucous glands		atrophic changes, mostly over the age of 70; decrease in the quantity and quality of secretions, leading to less hydrated vocal folds	reported in males; limited data on females	Caruso et al. (1995); Gracco and Kahane (1989); Sato and Hirano (1998); Schötz (2006); Tarafder, Datta, and Tariq (2012)
Innervation		some evidence of disruption/degeneration; decrease in neuromuscular control	no data on gender differences	Linville (2001); Caruso et al. (1995); Liss, Weismer, and Rosenbek (1990); Mueller (1997); Sataloff et al. (2017); Schötz (2006)
Superior laryngeal nerve		a decrease in both the size and number of myelinated fibers, leading to a diminution in laryngeal sensitivity	no data on gender differences	Kost and Sataloff (2020)
Vascular supply		reduced blood supply to the laryngeal muscles due to reduces blood vessel diameter and thickening of capillary walls	gender differences is unclear	Caruso et al. (1995); Linville (2001); Liss et al. (1990); Mueller (1997); Thomas et al. (2008)

According to M. Santos et al. (2021), presbylarynx is characterized by two or more of the following endoscopic/laryngoscopic findings: vocal fold bowing, prominence of vocal processes in abduction, and a spindle-shaped glottal gap. Besides these findings, other authors also reported vocal folds atrophy (Pessin, Tavares, Gramuglia, de Carvalho, & Martins, 2017; Pontes et al., 2005; Takano et al., 2010), a decrease or normal mucosal wave amplitude, a reduced phase of glottic closure and aperiodicity of glottis cycle (Galluzzi & Garavello, 2018).

From a vocal fold function perspective, males demonstrate an increased incidence of glottal gaps with aging, presumably as a consequence of vocal fold atrophy (Sataloff et al., 2017). In contrast, females do not differ in the overall incidence, but demonstrate different glottal gap configurations depending on age (Sataloff et al., 2017; Zamponi et al., 2021). Young women overwhelmingly demonstrate incomplete glottic closure and posterior glottal gaps, while older women have a higher incidence of anterior glottal gaps (Lenell et al., 2019; Linville, 1992).

Therefore, men and women experience opposite changes in the vocal fold structure and dynamics with aging, which may be explained by the sex-related senescence patterns of the endocrine system (Lenell et al., 2019; Zamponi et al., 2021). The presence of sex steroid hormone receptors within the vocal fold tissue suggests a relationship between hormones and vocal fold functions (Lã & Ardura, 2020; Lenell et al., 2019; Zamponi et al., 2021).

Although both males and females experience age-related endocrine function changes that result in declines in sex hormone production, the progressive decline of androgens in males (i.e., serum testosterone levels decrease by 1% per year after the age of 30 (Zamponi et al., 2021)) greatly differs from the dramatic and sudden decline of estrogen production following menopause in females (Lã & Ardura, 2020; Lenell et al., 2019). The age-related declines in testosterone, known as andropause, contributes to sarcopenia in aged males (over 65 years), which also affect the laryngeal muscles (Zamponi et al., 2021), and may contribute to age-related voice changes on older men. In contrast, menopause appears not to affect all women's laryngeal system similarly (Lenell et al., 2019). The serum concentrations of all three sex steroid hormones (i.e., estrogens, progesterone, and testosterone) drops, but at different rates (Lã & Ardura, 2020) leading to a relative increase of androgen levels in postmenopausal women (D'haeseleer et al., 2011; Lã & Ardura, 2020; Zamponi et al., 2021). These hormonal changes alter the physical characteristics of the vocal folds (i.e, increase of the vocal fold mass and drying of the laryngeal mucosa (Gorham-Rowan & Laures-Gore, 2006; Pontes et al., 2005, 2006)) and consequently affect the phonation process (Mautner, 2011). However, as adipose tissue is the primary source of estrogen production in postmenopausal women, some authors argue that reduced concentrations of estrogen are attenuate when body mass index values are high (Lã & Ardura, 2020; Lenell et al., 2019). Therefore, albeit the influence of hormones on females laryngeal senescence have been widely recognized, the mechanism of action of these hormones is scarcely understood (Sataloff et al., 2017).

2.2.3 Articulatory system

The final stage of speech production takes place in the supraglottal airway, called the vocal tract, which includes essential articulators for speech production (Linville, 2001; Mautner, 2011; Redford, 2015). The supraglottal system consists of the entire vocal tract above the glottis level, which includes the pharynx, the oral cavity and the nasal cavity (Linville, 2001; Makiyama & Hirano, 2017). In this sense, the movement of the articulators alters the shape of the vocal tract, which act as a resonator, to determine the phonetic quality of speech sounds (Redford, 2015). As a result, changes in the supraglottal system with age may also affect speech, as changes in articulatory positioning alter speech resonance characteristics (Galluzzi & Garavello, 2018; Linville, 1996; Schötz, 2006).

The basic contours of the pharynx can be fully developed by the age 9, but the vocal tract continues to grow in length and circumference through puberty and the full growth is usually not complete until age 20 (Sataloff et al., 2017). However, noticeable anatomical and physiological changes in the supralaryngeal system have been reported posteriorly, between young adulthood to old age (Sataloff et al., 2017). During this period, the craniofacial skeleton continues to grow (3%–5%) (Sataloff et al., 2017; Schötz, 2006). Additionally, a slight lowering of the larynx was observed, due to the atrophy of the head and neck muscles as well as the thinning of the intervertebral disks (more common in females), which leads to an increase of the length of the vocal tract (Boone, McFarlane, Berg, & Zraick, 2010; Linville, 2001; Linville & Fisher, 1985; Schötz, 2006).

In general, age-related changes in the vocal tract include: elasticity decrease and atrophy of the facial, the mastication and the pharyngeal muscles; a decrease of the soft-palate volume (Linville, 2001; Mahne et al., 2007; Massimo & Elisa, 2014; Sataloff et al., 2017; Schötz, 2006; Vipperla et al., 2010); enlargement of the the pharyngeal lumen (Hixon et al., 2018); and decline of the sensory innervation (Hixon et al., 2018; Martin et al., 1994). Extensive degenerative changes also occur in the temporomandibular joint (gradual reduction in size and reductions in blood supply), which controls the jaw movement during speech production (Sataloff et al., 2017; Schötz, 2006; Vipperla et al., 2010).

In relation to velopharyngeal-nasal changes with aging, studies have reported velopharyngeal muscles (tensor and levator veli palatini) and nasal mucosa atrophy (Kalmovich et al., 2005; Marino et al., 2018), a gradual increase of minimal cross-sectional areas and endonasal volumes (more prominent among women) probably due to decreased functioning of the nasal mucosa (Kalmovich et al., 2005; Marino et al., 2018). The nose tends to lengthen with age over the age of 65 (Bruzzi et al., 2017; Kalmovich et al., 2005; Sataloff, Johns, & Kost, 2015), due to changes in surrounding structures (i.e., midfacial loss of support) and/or in nasal cartilages (i.e., cartilages become thinner, atrophic and loose) (Kalmovich et al., 2005). However, the resistance of the nose and the airflow generally does not change across the lifespan (Sataloff et al., 2015).

In the oral cavity, a reduction of the oral tactile sensitivity (Bilodeau-Mercure & Tremblay, 2016) has been reported, as well as loss of lip strength (Klein, 1980; Linville, 2001; Mahne et al., 2007; Massimo & Elisa, 2014; Sataloff et al., 2017; Schötz, 2006; Vipperla et al., 2010; Wohlert & Smith, 1998) and maximal tongue strength (Bilodeau-Mercure & Tremblay, 2016) as function of age. The loss of elasticity and the thinning of the oral mucosa with the deterioration of attachments of the epithelium and connective tissue to bone are more apparent after age 70. However these changes may happen earlier in life (Klein, 1980; Sataloff et al., 2017; Schötz, 2006). Also, a diminished accuracy of the lower lip and jaw when performing rapid movements has been reported with age (Massimo & Elisa, 2014; Wohlert & Smith, 1998). Furthermore, the shape of the oral cavity may change with the loss of teeth and the introduction of dentures (Mautner, 2011), which may alter the occlusion and the articulation (Sataloff et al., 2017). In addition, in older adults there is often weight gain with an accumulation of fat in the neck and in the parapharyngeal space (Bruzzi et al., 2017).

The salivary glands undergo acinar atrophy, ductal irregularities, and parenchymal replacement by fibrous and/or adipose tissue with aging which causes a decline in salivary function (Klein, 1980; Mahne et al., 2007). The composition of saliva is also altered in older persons, with more viscous secretion (Aalami et al., 2003; Thompson & Chen, 2021). Diminished salivation can also be attributed to secondary factors such as medication (Aalami et al., 2003).

Regarding the tongue, age-related muscular changes include muscle atrophy (Bässler, 1987), decreased thickness of the lingual epithelium, reduced muscle fiber diameter, and fissuring of the tongue surface (Kuruvilla-Dugdale, Dietrich, McKinley, & Deroche, 2020; Sataloff et al., 2017), mainly over the age 50 (Thompson & Chen, 2021). Concerning the size of the tongue, contradictory findings have been reported in the literature (Goozée et al., 2005). While Sonies, Baum, and Shawker (1984) observed that the tongue of older people was reduced in thickness during rest, other studies have suggested that age-related changes in the tongue size are minimal to non-existent (Goozée et al., 2005; Kuruvilla-Dugdale et al., 2020; Mahne et al., 2007), which might be explained by the large increase of deposits of fatty tissue in the tongue to compensate the loss of muscle fibres (Bässler, 1987; Goozée et al., 2005; Kuruvilla-Dugdale et al., 2020). On the other hand, Klein (1980) suggested an increase in tongue size due to two factors: muscle tone loss and the need to expand to fill the oral cavity space as teeth are lost with aging. Chantaramanee et al. (2019) showed that the quality of the tongue muscles in healthy older speakers, assessed by the echo intensity of the tongue, is an indicator of tongue thickness at the middle and base of the tongue.

From a functional perspective, it has been reported that older individuals experience significant declines in tongue strength, although endurance remains stable throughout most part of life (Crow & Ship, 1996; Sataloff et al., 2017; Vanderwegen, Guns, Van Nuffelen, Elen, & De Bodt, 2013). Mortimore, Fiddes, Stephens, and Douglas (1999) and Mortimore, Bennett, and Douglas (2000) observed that the maximal tongue protrusion force decreases with age, while the tongue fatigability increases. Also, drying of the tongue and of the oral cavity may lead to an increase resistance of tongue movement across the palate during speech movements (Goozée et al., 2005).

Regarding the velopharyngeal function, there is less agreement among researches, with some studies indicating a higher nasalance for older adults (Hutchinson, Robinson, & Nerbonne, 1978; Marino et al., 2018; Styler, 2015), suggesting a decline in the control of velar movements as age increases (Bilodeau-Mercure & Tremblay, 2016; Hutchinson et al., 1978), and others suggesting no aged-related changes in nasalance, nasal air flow or in perceived nasality (Bilodeau-Mercure & Tremblay, 2016; Hixon et al., 2018; Hoit, Watson, Hixon, McMahon, & Johnson, 1994).

Despite speech production being a motor skill that is functional throughout the majority of life and being preserved even in the oldest of the old, the structures and function of the resonant organs undergoes significant age-related modifications (Bruzzi et al., 2017; Hixon et al., 2018). All of these changes affect the vowel production (Goozée et al., 2005; Vipperla et al., 2010), as changes in vocal tract dimensions and in tongue motion may influence the resonance patterns in older speakers (Vipperla et al., 2010; Xue & Hao, 2003).

2.2.4 Neuromuscular system

The production of speech sounds results from a complex sequence of neurologic, muscular, articulatory, and acoustic events (Raphael, Borden, & Harris, 2011). The sounds are produced by regulating the air flow, as it passes from the lungs to the exterior, by moving the vocal folds, the walls of the pharynx, the soft palate, the jaw, the tongue, and the lips to alter the shape of the vocal tract (Divenyi, Greenberg, & Meyer, 2006; Raphael et al., 2011). Thus, in order to produce intelligible and natural speech, the movements of the respiratory, laryngeal, and articulatory musculature have to be coordinated rapidly and with accurate timing (Divenyi et al., 2006; Hardcastle et al., 2012; Staiger, Schölderle, Brendel, & Ziegler, 2017). These movements are primarily the result of muscle contractions which are caused by nerve impulses, and the whole process is controlled by the nervous system (Raphael et al., 2011). In summary, the phono-articulatory function should be considered an integrated system (Bruzzi et al., 2017).

Age-related changes on motor function can be observed in both the peripheral and the central nervous system, with implications for speech production (Schötz, 2006; Vipperla, 2011; Waller, 2019). The decline of motor neurons in the peripheral neural system has been implicated as the main mechanism for muscle atrophy and loss of contractile strength in the muscles across older ages (Vandervoort, 2002; Vipperla, 2011). Other authors also reported that muscles in the older adults have fewer, but on average larger and slower motor units, which are able to partially compensate the losses, maintaining the contractile strength of affected muscles (Doherty, Vandervoort, & Brown, 1993; Linville, 2001). Additionally, older adults have demonstrated a decrease of the peripheral nerve conductivity that might be related with the decline of tactile and proprioceptive sensation with age (Aalami et al., 2003). A decrease in reflex time with age has also been reported (Aalami et al., 2003).

This age-related slowdown in the nerve conduction velocity might be behind the slower speech rate observed in older ages and might also disrupt coordination among articulators in older speakers (Hagen et al., 1996; Linville, 2001; Schötz, 2006; Vipperla, 2011). Previous studies about age effects in articulatory motor performance have demonstrated slower movements and reduced regularity in the rhythm of the tongue movements, reduced tongue retraction, and smaller reductions in distance traveled by the tongue during fast speech rates with age (Goozée et al., 2005; Hermes et al., 2018; Kuruvilla-Dugdale et al., 2020; Sonies, Baum, & Shawker, 1984).

In addition, the ability of adults to perceive pressure on the tongue also decreases with age (Caruso et al., 1995), as well as the tactile sensitivity of the lips diminishes as a function of age (Bilodeau-Mercure & Tremblay, 2016). This apparently reduced sensory ability is likely to negatively influence the motor performance associated with speech articulation. Reductions in (peripheral) nerve supply to muscles of the speech system might increase the muscle stiffness, leading to voice deterioration and breath support difficulties in the older adults (Linville, 2001). Additionally, degenerative changes in the superior laryngeal nerve might interfere with the regulation of the vocal folds vibration with aging (Linville, 2001).

Regarding the central nervous system, age-related changes in the brain structure include nerve cell losses (Linville, 2001), declines in brain volume and weight, demyelination, changes in cerebral vasculature (i.e., decrease of the cerebral blood flow and cerebral oxygen consumption mainly in areas with decreased gray and white matter), and changes in cortical-subcortical connectivity (Aalami et al., 2003; Dawson, 2020; Raz & Rodrigue, 2006). According to Raz and Rodrigue (2006), the anterior corpus callosum and the pre-frontal cortex are the most affected areas by these processes (Dawson, 2020; Raz & Rodrigue, 2006). Changes in the pre-frontal cortex may be the basis of declines in many executive functions that coordinate and control behavior (Dawson, 2020; Raz & Rodrigue, 2006). The age-related changes in the nerve cells of the cortex may also be responsible for the slowing down motor movements (Schötz, 2006). In addition, the age-related decline of dopamine levels and neurons in the cortex lead to deterioration of muscle tone, which impairs motor performance and reduces sensorimotor integration (Dawson, 2020; Linville, 2001; Waller, 2019).

Globally, these structural changes in the human brain may impact cognition in old age, as lesser volumes of the brain regions underling certain aspects of cognition have been associated to poorer performance in those domains (Dawson, 2020; Raz & Rodrigue, 2006). Therefore, neurotransmitter degeneration with age could slow speech rate in older speakers, but such slowing could also involve cognitive factors, such as deficits in linguistic processes, and decline in memory or attention function (Linville, 2001; Waller, 2019). However, the relationship between both aging and cognitive function in regard to speech production ability have been scarcely examined (Dawson, 2020).

This page intentionally contains only this sentence.

RELATED WORK

As reported in the previous section, the aging process causes specific alterations in the speech organs that are expected to play an important role in speech production. Despite the fact that the speech production appears to remain functional for conversation throughout the lifespan, even at older ages, most listeners can identify a speaker as an older adult when just hearing a few words (Hooper & Cralidis, 2009; Sataloff & Kost, 2020; Sataloff et al., 2017). In a same way, the speech acoustic changes (with aging) allow listeners to estimate the age of a person quite accurately (Rohloff, 2020; Schötz, 2006).

Following this line of thought, the human speech undergoes lifespan changes that may be identified through acoustic analysis (Spazzapan, Marino, Cardoso, Berti, & Fabbron, 2019), which provides objective measures that can be related to the subjective perceptual judgments (Mautner, 2011). Therefore, this is a non-invasive analysis, which offers quantitative information, enabling to infer indirectly the vibration patterns of the vocal folds, as well as the vocal tract (size and shape) in speakers of different sexes and age groups (Spazzapan et al., 2019). For these reasons, over the years, age-related variations on speech have been extensively investigated using acoustic analysis (Linville, 2001; Schötz, 2006; Vipperla et al., 2010).

Despite the literature on aging speech acoustics being extensive and diverse in terms of speech tasks and age groups assessed, aging speech is a complex process that remains poorly understood to date (Fuchs et al., 2021; Paulsen & Tillmann, 1998), namely its underlying articulatory adjustments (Fuchs, Gerstenberg, & Koenig, 2020; Xue & Hao, 2003). Therefore, there is a scarcity of research on the effects of aging on articulatory movements during speech production through direct measures (e.g., electropalatography (EPG), electromagnetic articulography (EMA), real-time magnetic resonance imaging (RTMRI), and US imaging).

In addition, as the speech signal is the mirror of the soul, the aging speech may reflect not only the physiological but also the psychological state of the speaker (Linville, 2001). In this vein, several studies have investigated the acoustic effects of diagnosed anxiety and depression (Cummins, Scherer, et al., 2015; Cummins, Sethu, Epps, Schnieder, & Krajewski, 2015; K. R. Scherer, 1986; S. Scherer, Morency, Gratch, & Pestian, 2015; Vicsi, Sztahó, & Kiss, 2012).

This chapter describes the previous acoustic and articulatory findings relevant to the understanding of the speech changes with age, and to the justification of the measures employed and the techniques chosen for investigation in this study. The impact of mood symptoms on speech acoustics is also reviewed.

3.1 Acoustic changes with age

The comprehension of the human communication process involves the analysis of speech in segmental and suprasegmental acoustic domains. The segmental acoustic features concern individual phonemes' characteristics (e.g., vowels). The suprasegmental or prosodic features refer to the phonetic and phonological studies of the relationships between syllabic units, without considering directly the segmental content (i.e., "what is said") but the sound form and its function ("how it is said") (Barbosa, 2019). The age-related changes may alter considerably the way individuals speak over time in both domains (Pellegrino, He, & Dellwo, 2018).

3.1.1 Acoustic vowel changes with age

The vowel formant measurements, F0 and vowel duration have a long history in the study of speech production (Kent & Vorperian, 2018), as the quality and intelligibility of each vowel can be described by these acoustic parameters (Adank, Van Hout, & Smits, 2004; Almurashi, Al-Tamimi, & Khattab, 2019; Kent & Vorperian, 2018; Maurer, 2016; Peterson & Barney, 1952; Themistocleous, 2017; Tykalova et al., 2020; C. I. Watson & Harrington, 1999; Zahorian & Jagharghi, 1993).

The vowel formant measurements are suited to articulatory interpretations of acoustic data (i.e., formant frequencies reflect the length and configuration of the vocal tract) (Das, Mandal, Mitra, & Basu, 2013; Eichhorn et al., 2018; Fant, 1970; Kent & Vorperian, 2018; McDougall & Nolan, 2007). As the first and second formant frequencies (F1 and F2) have a well-defined acoustic-articulatory relationship (reflecting primarily tongue position and lips rounding), they were used for the definition of several derived metrics such as: vowel space area (VSA), formant centralization ratio (FCR), vowel articulation index (VAI), and formant range ratios (i.e., F1RR and F2RR). F0 is determined by the rate of vibration of the vocal folds (Ladefoged & Johnson, 2011). Moreover, F0 is related to the anatomophysiological characteristics of the vocal folds, namely length, mass, vibration, tension and stiffness during phonation (Spazzapan et al., 2019; Titze, 1994). Some studies have recognized that temporal information has also been important for characterizing vowel quality (Almurashi et al., 2019; C. I. Watson & Harrington, 1999; Williams & Escudero, 2014), as the duration of speech segments can affect the spectral quality of phonemes (Fletcher, McAuliffe, Lansford, & Liss, 2015).

Despite the fact that numerous studies have evaluated the effects of aging on the acoustic properties of vowel segments (mainly for English) (Albuquerque et al., 2014; Eichhorn et al., 2018; Fletcher et al., 2015; Fougeron, Guitard-Ivent, & Delvaux, 2021; Goy, Fernandes, Pichora-Fuller, & van Lieshout, 2013; Jacewicz, Fox, & Salmons, 2011a, 2011b; Km & Yoon, 2012; Kyriaki, 2021; Linville & Fisher, 1985; Pellegrini et al., 2013; Ramig & Ringel, 1983; Schötz, 2006; Tykalova et al., 2020; Vipperla et al., 2010; Xue & Hao, 2003), most of them have focused solely on sustained vowel production (Dehqan, Scherer, Dashti, Ansari-Moghaddam, & Fanaie, 2013; Linville & Fisher, 1985; Rastatter & Jacques, 1990; I. R. Santos, 2005; Sebastian, Babu, Oommen, & Ballraj, 2012; Spazzapan et al., 2020; Stathopoulos, Huber, & Sussman, 2011; Xue & Deliyski, 2001; Xue, Jiang, Lin, Glassenberg, & Mueller, 1999). Thus, the primary purpose of the following subsections is to review the age-related acoustic changes on vowel segments in connected speech. Specifically, focus on the effects of age in the acoustic features F0, formant frequencies (F1, F2 and F3) and vowel duration. Notwithstanding, note that the comparison of results reported in the literature is not straightforward due to methodological differences among studies (various languages, different age spans, but also variable speech material and speech tasks) (Fougeron et al., 2021).

Fundamental Frequency

F0 has been found to be sensitive to the aging effect (Mautner, 2011). In females, the literature is highly consistent with the decrease of F0 with advancing age (Eichhorn et al., 2018; Higgins & Saxman, 1991; Kaur & Narang, 2015; Mautner, 2011; Mifune, Justino, Camargo, & Gregio, 2007; Schötz, 2006; Torre III & Barlow, 2009; Tykalova et al., 2020; Winkler, 2004). For instance, Schötz (2006) and Tykalova et al. (2020) have reported that female F0 decreased until the age 50 or 60 and then remained relatively stable. In Eichhorn et al. (2018) the young adult women ([20-30]) had significantly higher F0 than middle-aged ([40-60]) and older ([70-92]) women for all vowels. Nonetheless, there were no significant F0 differences between middle-aged women and older women, despite a slightly increase in the last age (Eichhorn et al., 2018).

Concerning males, there is no general agreement on the trend of changes in F0 due to aging. While Vipperla et al. (2010) indicated that the F0 reduces significantly above 60 years of age (i.e., F0 for older males [60-85] is about 15 Hz (10%) lower than for adult males [30-45]), other studies suggested that F0 in males decreases until age 50 and increases again at an advanced age (Schötz, 2006; Torre III & Barlow, 2009). On the other hand, Mautner (2011) reported that F0 increases continuously between the ages 42 and 93, more sharply between the ages [60-69] and [70-79], mainly in the vowel /i/. Additionally, most of the literature has reported no significant differences with age in males, with some authors reporting a slightly increase (Eichhorn et al., 2018; Higgins & Saxman, 1991; Winkler, 2004), and others indicating a slightly decrease of F0 (Mifune et al., 2007). Tykalova et al. (2020) reported no age-related changes in F0 as age increases (between the ages 20 and 90) but observed a slightly increase after the age 50. As the majority of the studies compare solely two age groups (e.g., young versus older or middle-aged versus older) and tend to consider different age ranges, it is hard at this time to draw solid conclusions about the non-linear effects of age on males' F0, and also to identify the age ranges in which changes occur.

The majority of aforementioned studies were cross-sectional in nature; yet, longitudinal studies of vowel segments have reported lowering of F0 for both women (Harrington, Palethorpe, & Watson, 2007; Mwangi et al., 2009; Reubold, Harrington, & Kleber, 2010) and men (Harrington et al., 2007). Nonetheless, for males there is also less agreement among researchers, with some longitudinal studies indicating a falling-rising pattern in F0 with age (Reubold & Harrington, 2015; Reubold et al., 2010).

Regarding EP, the few data available indicated an F0 drop of 20 Hz with advanced age in women, while for men no age-related changes were observed, when comparing young adults (aged 19-30) with two groups of older adults (aged 60-75 and over 75) (Pellegrini et al., 2013). Another study (Albuquerque et al., 2014), with speakers aged between 60 and 90, indicated a slight age-related F0 decrease in women and a trend of increase in men.

F0 changes with age during adulthood probably due to the effect that the normal course of physical aging has on individuals (Mautner, 2011). Thus, as F0 is related to the characteristics of the vocal folds and their structure changes differently with aging by gender, age-related gender differences on F0 are not unexpected. The literature has suggested that the F0 drop in women with age results from the increase in vocal fold mass due to hormonal changes that occur during menopause (Ferrand, 2002; Higgins & Saxman, 1991; Linville, 2001; Ma & Love, 2010; Mautner, 2011; Pontes et al., 2005;

Sataloff, Caputo Rosen, Hawkshaw, & Spiegel, 1997). The raise in F0 in men after middle age has been attributed to reduced vocal fold mass and stiffness of vocal fold tissues due to aging-induced atrophy of the internal thyroarytenoid (Higgins & Saxman, 1991; Hollien & Shipp, 1972; Linville, 2001; Nishio & Niimi, 2008; Sebastian et al., 2012).

Formant frequencies

As the physical constitution of the vocal tract changes during the process of aging, the formant frequencies are expected to change as well (Steffens, 2011). Namely, the lowering of the larynx with age alters the volume of the vocal tract, which may lead to a decrease of the vowel formants (Kyriaki, 2021; Torre III & Barlow, 2009; Xue & Hao, 2003). Nonetheless, the conclusions across studies are inconsistent, with some studies showing an age-dependent formant frequency lowering (especially F1) in both genders (Hazan et al., 2018; Kaur & Narang, 2015; Linville, 2001; P. J. Watson & Munson, 2007; Xue & Hao, 2003), and others reporting no changes in formant frequencies (Benjamin, 1982; Fletcher et al., 2015), mostly in higher formant frequencies (F3 and F4) (Eichhorn et al., 2018; Leung, Oates, Papp, & Chan, 2020; Tykalova et al., 2020). Beyond that, there are vowels that have presented a different variation pattern of the formant frequencies with age and gender (Eichhorn et al., 2018; Mautner, 2011; Schötz, 2006; Scukanec, Petrosino, & Squibb, 1991; Torre III & Barlow, 2009; C. I. Watson & Hui, 2010).

Variations in F1 and F2 can lead to changes in the dimension of the vowel space, which can be measured through different metrics. Some studies reported a centralization of the vowel space in older speakers (which should result in movement to the centroid of formant space) (Kyriaki, 2021; Liss et al., 1990; Mertens, Mücke, & Hermes, 2020; Rastatter, McGuire, Kalinowski, & Stuart, 1997; Schötz, 2006; Scukanec et al., 1991; Torre III & Barlow, 2009). Meanwhile, Tykalova et al. (2020) and Fougeron et al. (2021) reported a tendency for VSA to increase between the ages 20 and 90/93 in both genders, which may be related, partly, to longer vowel duration in the older ages (Tykalova et al., 2020).

Despite the lack of literature about age-related changes using other vowel space metrics, some authors have reported no changes in the FCR for both genders (Fougeron et al., 2021); increase of the F2RR in males after their 40's (Fougeron et al., 2021); decrease of the F1RR until the age 57 (Fougeron et al., 2021); and decrease of the VAI with age in both genders (Mertens et al., 2020).

Most of the longitudinal studies (all with English speakers) have shown a decrease in F1 with aging in women (Harrington et al., 2007; Mwangi et al., 2009; Reubold et al., 2010); for men there is less agreement across researches, with some studies indicating a falling-rising pattern in F1 (Reubold & Harrington, 2015; Reubold et al., 2010), and others suggesting a lower F1 as age increases (Harrington et al., 2007). Regarding vowel space, Gahl and Baayen (2019) reported a vowel space expansion throughout young to middle-aged adulthood (ages 21 - 49), regardless of the vowel duration.

Until now the majority of the literature on age-related changes in vowel formants concerns studies with English speakers. Only Mertens et al. (2020), Tykalova et al. (2020), Fougeron et al. (2021), and Kyriaki (2021) report speech data from adults who speak a different language (i.e., German, Czech, French and Greek, respectively). On what concerns EP, the available data have shown that age-related changes in vowel formant frequencies are not consistent and seem to be different among vowels (Albuquerque et al., 2014; Pellegrini et al., 2013; Rato et al., 2017). Pellegrini et al. (2013) showed a greater centralization of the vowel space in younger speakers (males and females aged [19-30] versus [60-90]), whereas in Albuquerque et al. (2014) a sightly decrease of F1 and F2 with aging was observed, mainly in females. Still, in Albuquerque et al. (2014), comparing the data obtained for older adults between the ages 60 to 90, with the data of a previous study with young EP adult speakers (Escudero et al., 2009), a trend towards the centralization of vowels with aging was observed. Rato et al. (2017) studied the vowel space of young ([20-30]) and middle-aged ([50-60]) speakers of Braga, which were recorded at the same time slot, and reported no significant age-related changes. On the other hand, comparing young adults recorded in 1996-97 with middle-aged speakers recorded in 2012-13, Rato et al. (2017) observed a wider vowel space in older speakers, which seems to be more related with gradual language changes in the speech community of Braga than with age-related vocal tract modifications.

Taking into account the inconclusive results on age-related formant frequencies changes, Linville (2001) suggested a "mixed" or "blended" model of vocal tract resonance changes with aging in which an interaction exists between gender, the resonance effects of vocal tract lengthening, and vowel articulatory patterns.

Despite the fact that in speech production the articulators move continuously in space and time (Morrison & Assmann, 2013) (i.e., speech is intrinsically dynamic (Yuan, 2013)), vowels have been traditionally characterized by static cues such as F1 and F2 obtained at vowel midpoint (e.g., Kent and Vorperian (2018); Peterson and Barney (1952)). Nevertheless, the explanations that have been advanced to account for age-related changes in vowel formant frequencies have referred alterations in speakers' anatomical dimensions and also in movement of the individual's speech organs (Das et al., 2013). Therefore, a dynamic approach to study the aging effect on vowel production could be important to account for specific articulatory adjustments during speech (Eichhorn et al., 2018; Linville, 2001), since a static approach only reflects anatomical differences among speakers (Gahl & Baayen, 2019). However, the effect of aging on dynamic properties of vowels has not been studied thoroughly. The cross-generational and cross-dialectal studies of Jacewicz et al. (2011a, 2011b) revealed substantial differences both in formant dynamics and vowel dispersion in the acoustic space according to age for three American English dialects.

Duration

Delving into age-related differences in the vowels duration, several studies have reported that they are longer in older speakers than in younger adults, not only in English (Benjamin, 1982; Fletcher et al., 2015; Liss et al., 1990; Morris, 1986; Smith, Wasowicz, & Preston, 1987), but also in Swedish (Schötz, 2006), in EP (Albuquerque et al., 2014; Pellegrini et al., 2013), in German (Mertens et al., 2020), in Greek (Kyriaki, 2021), in French (D'Alessandro & Fougeron, 2018), and in Czech (Tykalova et al., 2020; Volín, Tykalová, & Bořil, 2017). Studies have suggested that the vowels duration is lengthened mainly over the age of 50 (D'Alessandro & Fougeron, 2018; Kyriaki, 2021; Schötz, 2006; Volín et al., 2017). Smith et al. (1987) reported that segment durations (both consonants and vowels) in older speakers (males and females between ages 66 to 75) are 20%-25% longer than in younger adults (24 to 27 years of age). In spite that Volín et al. (2017), when analyzing reading speech data of 200 Czech speakers between the ages 20 and 80, observed that the duration of female vowels follows a near-linear trend of increase with age; whereas the duration of male vowels shows a v-shape trend, decreasing until middle-aged and increasing sharply thereafter. Schötz (2006) also reported that stressed vowel duration increases significantly with age, but in two specific vowels ([a] and [iː]) a v-shape trend was observed.

Despite the fact that the increase of vowel duration has been one of the most robust age-related effects set up in the literature, some studies have not found statistically significant differences in segments/vowels duration with age for both genders (Jacewicz et al., 2011b; Mautner, 2011). Linville (2001) suggested that older adults show some flexibility in segment durations. That is, in a speaking situation that requires longer utterances, older adults frequently reduce the phoneme durations to maintain the physiological support needed to successfully produce the utterance (Linville, 2001).

The reason why the vowels duration is extended as speakers age has not been fully explained. As the segment duration depends, in part, on the speech rate, which seems to decrease in older adults (Linville, 2001; Schötz, 2006; Smith et al., 1987), a slower speech rate results in an increase of the segments duration with aging. This probably occurs as a consequence of the lessening in the nerve conduction velocities, respiratory changes, increased cautiousness and to the adjustment by older speakers of their tempo to maintain speech fluency (Linville, 2001; Schötz, 2006).

3.1.2 Suprasegmental changes with age

From a acoustic point of view, the human speech cannot be fully characterized as the manifestation of sequences of phonemes, syllables or words (Nooteboom, 1997). In this sense, speech properties that cannot be obtained from the underlying phonemes are often called suprasegmental characteristics of speech (Nooteboom, 1997). Thus, the suprasegmental features comprise the physical correlates of F0 modulation (i.e., intonation), intensity, temporal patterns (i.e., rhythm) and also voice quality (Barbosa, 2019; Crystal, 1969; Nooteboom, 1997; Prieto & Roseano, 2018).

Most studies have focused on the production of segments, such as vowels, the suprasegmental features of speech are also particularly vulnerable to age (Pellegrino et al., 2018). However, the effect of aging on suprasegmental characteristics of speech has been investigated in far less detail, namely in spontaneous speech.

Speech rhythm

Temporal aspects of speech, as part of the speech rhythm (Nooteboom, 1997), seem to be strongly affected by the age of the speaker (Fougeron et al., 2021; Schötz, 2006), namely the rate of speech. This is a physical, acoustic, and physiological phenomenon which is connected to the length of the segmental units and the number of the pauses that appear between the rhythmical groups made of segmental and suprasegmental elements (Michalik, Kaczorowska-Bray, Milewski, & Solak, 2018). In general, despite the fact that articulation rate corresponds to the number of syllables produced in the speech sample within a certain period of time (without pauses), and the speech rate includes the duration of the pauses that were made within the speech sample (Michalik et al., 2018), some previous studies do not clearly specify the metric used. As a consequence of this methodological difference (and also others such as variable speech material/tasks and different age spans), the comparison of the results reported in the literature is not straightforward.

Nevertheless, there have been various studies that show evidence for a slowing of both the speech rate (Bóna, 2014; Fougeron et al., 2021; Goy, Kathleen Pichora-Fuller, & van Lieshout, 2016; Huber & Spruill III, 2008; Kyriaki, 2021) and the articulation rate (Bóna, 2014; Bourbon & Hermes, 2020; Fougeron et al., 2021; Hazan et al., 2018; Hermes, Bourbon, & Cecile, 2020; Jacewicz, Fox, O'Neill, & Salmons, 2009; Jacewicz, Fox, & Wei, 2010; Kyriaki, 2021; Mertens et al., 2020; Pellegrini et al., 2013; Schötz, 2006; Tykalova et al., 2020; Volín et al., 2017) with increasing age, regardless of the speaker's language or the speech type (i.e., reading speech, spontaneous speech or conversation). However, when the articulation rate in both reading and spontaneous speech was investigated, the results were somewhat contradictory: Brückl and Sendlmeier (2003) found a decrease of the reading rate in German female speakers with increasing age, whereas in articulation rate of spontaneous speech (i.e., a picture description task) no significant change was observed. Jacewicz et al. (2010) studied 150 adults of both genders between the ages 20 and 91 and reported that articulation rate in reading decreases after the age [20-34], while for articulation rate in spontaneous speech, adults in their 40s had the higher values (i.e., articulation rate only decreases after this age). Volín et al. (2017), for Czech population, also observed a deceleration in the articulation rate (in reading speech) as a function of age, but the trends seem not to be linear (i.e, middle-aged men are the fastest).

Regarding longitudinal studies, a decrease in speech and articulation rate with advancing age was also observed for an Italian journalist, Piero Angela (in a corpus of reading speech collected at the ages of 40 and 79) (Massimo & Elisa, 2014). Another longitudinal study, based on five public talks produced by Noam Chomsky between the ages 40 and 89, also detected a statistically significant decrease in speech rate measures (Pellegrino, 2019). In Queen Beatrix's formal speeches (between ages 42 and 74) a decrease in articulation rate in the first decades and an increase in the last decade was observed (Quené, 2013). Additionally, Gerstenberg, Fuchs, Kairet, Frankenberg, and Schröder (2018) analyzed the articulation rate of five French and five German speakers, using interviews in two moments with ten years of difference (between the 70s and the 80s), and reported a decrease in the articulation rate for German participants and an increase for French speakers (Gerstenberg et al., 2018). For EP, Valente, Oliveira, Albuquerque, Teixeira, and Barbosa (2021) studied three interviews of a famous native speaker, and observed a significant decrease in speech rate with age, namely between the first age (51) and the last ages (i.e, 74 and 82). This tendency of the decrease for rhythmic parameters with age was also observed in a submitted longitudinal study, by the same authors, with three famous speakers.

Although the majority of the studies point to a decrease in speech and articulation rate with age, the relationship between age and rate of speech is unclear (Fletcher et al., 2015). Are the limitations in neuromuscular speed directly responsible for the speakers' reduced articulation rate or is a slower rate of speech and longer vowel durations compensatory mechanisms for maintaining the articulatory precision (Fletcher et al., 2015)?

Another aspect of the aging speech is the speech pauses. Several studies have shown that duration as well as the use of speech pauses tend to increase with aging (Dimitrova, Andreeva, Gabriel, & Grünke, 2018; Fougeron et al., 2021; Hartman & Danhauer, 1976; Hermes et al., 2020; Massimo & Elisa, 2014; Pellegrini et al., 2013; Steffens, 2011). In contrast, Hazan et al. (2018) observed a significant decrease in the frequency of pauses in the females' spontaneous speech/conversation (i.e., produced during a problem-solving task (diapix)), and no differences in males.

Bóna (2014) reported an increase in the frequency of pauses in different speech tasks, while no changes were observed in the ratio of pauses and in the duration of pauses with age regardless of the speech task (e.g., reading aloud, spontaneous speech, narrative recalls, or conversation). Similarly, for Brazilian Portuguese, V. Martins and Andrade (2011) observed a significant increase in the frequency of pauses in spontaneous speech (picture description task) with age, namely between the ages [60-79] and ≥ 80, even though no changes in mean pause duration or in the percent pause time occurred with age.

The use of more speech pauses, and the associated slowing of speech rate can directly be traced to an aggravation of the lung-function (Schötz, 2006; Steffens, 2011). In addition, pauses may give time for older speakers to solve difficulties of speech planning (e.g., lexical access) and articulation (Bóna, 2014).

Regarding the speech intervals, Huber and Spruill III (2008) observed that young adults produced significantly longer utterances than older adults in spontaneous speech. Thus, the duration of the speech intervals tends to decrease with aging (Hoit & Hixon, 1987; Huber & Spruill III, 2008). Producing shorter utterances allows older adults to take more frequent breaths, which may be a compensatory mechanism for changes in respiratory physiology, but more frequent breath pauses may have unintentional linguistic consequences (Redford, 2015).

Intonation - Speaking F0

From a production point of view intonation is related to the rate of vibration of the vocal folds and its main correlate in the acoustic dimension is the F0 (Madureira, 2016), which in connected speech is often referred to as speaking F0 (Leung, Oates, Papp, & Chan, 2022).

Similarly to vowels' F0, speaking F0 has been found to be sensitive to aging, tending to decrease in older females, in both reading (Benjamin, 1981; de Pinto & Hollien, 1982; Dimitrova et al., 2018; Fougeron et al., 2021; Goy et al., 2013, 2016; Leung et al., 2022; Morgan & Rastatter, 1986; Murry, Brown, & Morris, 1995; Nishio & Niimi, 2008; Russell, Penny, & Pemberton, 1995; P. T. Silva et al., 2011; Spazzapan et al., 2018; Stoicheff, 1981; Volín et al., 2017) and spontaneous speech (Hazan et al., 2018; Murry et al., 1995; Winkler, 2004). Fougeron et al. (2021) observed that speaking F0 decreases until the age of 40 and then remains stable, while Pegoraro-Krook (1988) reported a decrease until the age [60-79] and then a slight increase. Additionally, Berg et al. (2017) did not find changes in females' speaking F0 in spontaneous speech (automatic speech - count from 21 to 30), but only analyzed females between the ages 40 and 79.

For men, most studies have reported that speaking F0 decreases from young adulthood into middle age and then rises again into older ages (Fougeron et al., 2021; Hollien & Shipp, 1972; Morris, Brown Jr, Hicks, & Howell, 1995; Pegoraro-Krook, 1988; Ramig & Ringel, 1983). Other studies have indicated that speaking F0 decreases (Benjamin, 1981; Cox & Selent, 2015; Decoster, 2000; Spazzapan et al., 2018), while others have suggested a speaking F0 increase (or slight increase) in older age (Berg et al., 2017; Dunashova, 2021; Volín et al., 2017; Winkler, 2004). Additionally, other studies have reported no speaking F0 changes in males (Goy et al., 2013, 2016; Hazan et al., 2018; Leung et al., 2022; Murry et al., 1995). Note that, some of these contradictory findings may occur due to methodological differences between studies, such as the comparison of different age ranges. For example, Spazzapan et al. (2018) reported a speaking F0 decrease after the age 40, but only analyzed reading speech of speakers between the ages 19 to 59; while Fouquet, Pisanski, Mathevon, and Reby (2016) in a longitudinal study with males, between the age 7 and 56 (interviews), observed a speaking F0 stability between the ages 21 and 56.

Regarding F0 modulation or speaking F0 variability, limited research has been carried out and the results have not been consistent. Some studies point to an increase (Benjamin, 1981; Berg et al., 2017; Decoster, 2000; Morris & Brown, 1994; Morris et al., 1995; Stoicheff, 1981; Volín et al., 2017) and others to a decrease (Goy et al., 2013; Leung et al., 2022) or no changes (Berg et al., 2017; Goy et al., 2016; Leung et al., 2022) of the speaking variability with aging. Benjamin (1981) observed that older speakers (males and females) produced larger intonational ranges, and more inflections in reading than younger ones. Dimitrova et al. (2018) analyzed reading speech of females and reported that older speakers ([79-88] years) used a wider pitch range than the younger speakers ([19-23] years). Fougeron et al. (2021) only confirmed this trend for females (no changes with aging in males), reporting that older women (after the age 75) exhibited more pitch modulation than younger ones. This increase in modulation may result from a decreased stability in the control of F0 (as for sdF0 on sustained vowels), but also from a different prosodic organization of the sentence inducing more pitch accents (Fougeron et al., 2021). Furthermore, Markó and Bóna (2010) also concluded that age affects the F0 range between speech types in opposite ways: reading was produced with a somewhat wider F0 range than spontaneous speech in younger speakers, whereas F0 range was wider for spontaneous speech than for reading in old speakers.

For EP, previous studies with different speech corpora (e.g., reading or conversation), indicated a slight or significant age-related speaking F0 decrease in women, and a non-agreement in men (Guimarães & Abberton, 2005; Pellegrini et al., 2013). In a longitudinal study with a famous EP native speaker, the speaking F0 and the standard deviation of speaking F0 presented an inverted v-shape trend, with a significant increase between the age 51 and 74 and a significant decrease between age 74 and 82 (Valente et al., 2021). Comparing only the younger and the older age, these features presented a tendency for a significant increase, meaning a higher F0 mean, more F0 variability and higher F0 peaks with aging (Valente et al., 2021).

Intensity

Studies about age-related changes on speech intensity are scarce, and generally they have focused on sustained vowels. In connected speech the majority of the literature reported no significant changes with aging on speech intensity (Goy et al., 2013; Huber, 2008; Huber & Spruill III, 2008; Morris, 1986; Morris et al., 1995; Pellegrino, He, & Volker, 2021). Only Ryan (1972), Berg et al. (2017) and

Kyriaki (2021) have reported a significant increase, whereas Goy et al. (2013) and Schötz (2006) have observed a significant decrease of the mean speech intensity, mainly in males, as age increases.

While age-related decrease in subglottal air pressure and glottal competence may be linked to a reduced vocal intensity (Eadie, 2000; Kyriaki, 2021; Linville, 2001; Ramig et al., 2001; Schultz, Rojas, St John, Kefalianos, & Vogel, 2021), the age-related decline in the hearing level could also be responsible for the higher speech intensity observed in the older ages (Kyriaki, 2021).

Concerning speech intensity variation, Volín et al. (2017) and Schötz (2006) have indicated an increase, while Pellegrino et al. (2021) observed a decrease as a function of age. A longitudinal study over a period of 5 years on 11 healthy male speakers (age ranging from 50 to 81) only revealed a slight (non significant) decrease in intensity range (Verdonck-de Leeuw & Mahieu, 2004).

Voice quality

Many studies have been quoted in the literature regarding changes in voice quality with aging (Fougeron et al., 2021; Jayakumar, Benoy, & Yasin, 2020; Spazzapan et al., 2019). In sustained vowels, noise measurements (i.e., harmonic-to-noise-ratio (HNR) or noise-to-harmonic ratio (NHR)), and measures of F0 and amplitude stability, namely jitter, shimmer, F0 standard deviation and amplitude standard deviation, are frequently used to describe the changes in the vocal characteristics across age (Linville, 2001; Spazzapan et al., 2019). Generally, a decline in voice quality (e.g., increase instability of the speech signal) has been reported in older adults (Fougeron et al., 2021; Jayakumar et al., 2020), and has been linked to the perceptual correlates of increased hoarseness, harshness, strain and/or vocal tremor (Linville, 2001). This increased amount of noise and instability in the speech signal has been attributed to differences in anatomical, physiologial, and neural control with aging (Linville, 2001; Spazzapan et al., 2019). However, as for speech intensity measures, studies about age-related changes on voice quality using connected speech samples are scarce (Spazzapan et al., 2019).

To the best of our knowledge, only Lortie, Thibeault, Guitton, and Tremblay (2015), Dunashova (2021) and Schötz (2006) have reported HNR data obtained in connected speech samples. Lortie et al. (2015), for native speakers of Canadian French, observed a lower HNR in connected speech compared with sustained vowels, and also a decrease in HNR with age, which is consistent with previous studies using sustained vowels (Ferrand, 2002). Nonetheless, for Swedish almost no age-related variation was found in females' HNR, while in males, a slight decrease followed by an increase after age 50 was observed (Schötz, 2006). The longitudinal study of Dunashova (2021), using reading samples of one male speaker at different ages (i.e., David Crystal's speeches at age 59 and 74), also found no changes in HNR with age.

3.2 Impact of mood symptoms on speech

The age-related changes on speech have been shown to negatively impact the communication and the quality of life (i.e., aging speech can have a negative impact on independence, integration, and effective communication) (Desjardins et al., 2022; Lortie et al., 2015; Selent, 2014; Stathopoulos et al., 2011). Indeed, older adults have reported difficulties to speak in some situations, such as in a noisy condition or on the telephone, which can cause anxiety, frustration, and even social isolation (Lortie et al., 2015; Roy et al., 2007; Schneider, Plank, Eysholdt, Schützenberger, & Rosanowski, 2011; Selent, 2014; Verdonck-de Leeuw & Mahieu, 2004).

Psychological disorders, such as depression and anxiety, are major public health concerns defined by a combination of atypical perceptions, behaviors, thoughts, emotions and relationships with others (World Health Organization, 2018). Despite that there is some evidence that aging is associated with an intrinsic reduction in susceptibility to anxiety and depression (Jorm, 2000), World Health Organization recognizes that depression and anxiety are not rare psychological conditions in the older population, which affect approximately 7% and 3.8% of the world population aged 60 and over, respectively (World Health Organization, 2017a). The acoustic parameters of speech are sensitive to the psychological and physiological changes of anxiety and depression (K. R. Scherer, 1986; S. Scherer et al., 2015; Vicsi et al., 2012). In this sense, speech, as a biomarker, may contribute and assist specialists in more accurate and objective detection of anxiety and depression symptoms (Cummins, Scherer, et al., 2015; Taguchi et al., 2018).

3.2.1 Anxiety symptoms and acoustic features

Fear, tension and distress are common symptoms associated with anxiety, usually assessed by subjective methods (Laukka et al., 2008; Özseven, Düğenci, Doruk, & Kahraman, 2018). As anxiety has a reflection in people's voice due to the somatic symptoms associated with the respiratory system, the acoustic parameters could be used as an objective method to assist in the assessment of anxiety symptoms (Özseven et al., 2018; Sataloff, 2005). Several research studies have evidenced the influence of anxiety symptoms in acoustic parameters. According to some authors, the mean F0 increases in individuals with anxiety (Banse & Scherer, 1996; Diamond, Rochman, & Amir, 2010; Hagenaars & van Minnen, 2005; Low, Bentley, & Ghosh, 2020; Weeks et al., 2012). The F0 variability also demonstrated to be a good indicator of anxiety symptoms, as Hagenaars and van Minnen (2005) and Goberman, Hughes, and Haydock (2011) have reported a superior pitch variability with the increase of anxiety. Other studies, in contrast, found different trends in this acoustic feature (Drioli, Tisato, Cosi, & Tesser, 2003; Protopapas & Lieberman, 1997; Ververidis & Kotropoulos, 2006).

Increased anxiety also leads to higher jitter and shimmer values (Fuller, Horii, & Conner, 1992; Low et al., 2020). Loudness and HNR, on the other hand, have an irregular performance, presenting distinct results in different research studies - either no change, decrease or increase (Murray & Arnott, 1993; Siegman & Boyle, 1993; Wörtwein, Morency, & Scherer, 2015).

Özseven et al. (2018) analyzed a broader set of acoustic parameters (122 acoustic measures) in patients diagnosed with anxiety and in healthy individuals and observed significant changes in 42 acoustic parameters (e.g., F0, F1, shimmer, jitter, mel-frequency cepstrum coefficients (MFCCs), and wavelet coefficient), with different directions and intensities, with anxiety. For example, F0 mean, F1 mean, shimmer, jitter and wavelet coefficients increase in anxiety patients and, in general, MFCCs decrease with anxiety.

Finally, the suprasegmental measures, number of pauses and percent pause time are positively correlated with the increase of anxiety (Goberman et al., 2011; Laukka et al., 2008; Wörtwein et al., 2015). However, speech rate tends to increase with the increase of anxiety (Hagenaars & van Minnen, 2005; Laukka et al., 2008; Murray & Arnott, 1993).

3.2.2 Depression symptoms and acoustic features

Depression causes changes in the somatic and autonomic nervous system, which is reflected in muscle tension and respiratory rate (Ellgring & Scherer, 1996; Won & Kim, 2016). These changes have an impact on prosody and speech quality (Alpert, Pouget, & Silva, 2001; S. Scherer et al., 2013; Yang, Fairbairn, & Cohn, 2013). The increased muscle tension and the changes in salivation and the mucus secretion affect vocal tract and limit articulatory movements, leading to articulation errors, reduced pitch range, a decrease in speech rate and increased hesitations (Cannizzaro, Harel, Reilly, Chappell, & Snyder, 2004; Ellgring & Scherer, 1996). In a vast amount of studies, the reduction of both F0 range and the F0 average are found to be linked with depression severity (Breznitz, 1992; Hönig, Batliner, Nöth, Schnieder, & Krajewski, 2014; Kuny & Stassen, 1993; Low et al., 2020; Mundt, Snyder, Cannizzaro, Chappie, & Geralts, 2007). F0 range was also evidenced to be a biomarker in treatment responders, as pitch variability increases significantly in patients who present a decrease in depressive symptoms (Mundt et al., 2007; Mundt, Vogel, Feltner, & Lenderking, 2012).

The slowing of thoughts and reduction of physical movements that occurs in depression (i.e., psychomotor retardation) could explain the reduction in F0 parameters, as the complexity of the larynx neuromuscular system is affected by disturbances in muscle tension due to psychomotor retardation (Bennabi, Vandel, Papaxanthis, Pozzo, & Haffen, 2013; Cannizzaro et al., 2004; Greden, 1993; Roy, Nissen, Dromey, & Sapir, 2009; Sobin & Sackeim, 1997). The increased muscle tension in the vocal tract could also explain the tightening of the vocal folds and, consequently, less variable speech (Cannizzaro et al., 2004; Ellgring & Scherer, 1996; Horwitz et al., 2013; Nilsonne, 1987; Quatieri & Malyska, 2012; Sobin & Sackeim, 1997). Nevertheless, other studies did not find a significant correlation between the F0 parameters of depressed and non-depressed patients, possibly due to methodological aspects or the intrinsic characteristics of F0 (i.e., simultaneously an indicator of the affective status and a marker of the physical state of vocal folds) (Cannizzaro et al., 2004; Mundt et al., 2007; Quatieri & Malyska, 2012).

Similarly to F0, contradictory results concerning variation in loudness were found in the literature, showing either significant and non-significant improvement of energy parameters after depression treatment (Alpert et al., 2001; Kuny & Stassen, 1993; Stassen, Bomben, & Günther, 1991).

Considering that depression could affect vocal tract properties, formant features are also suitable as a marker of these changes (Cummins, Scherer, et al., 2015; Flint, Black, Campbell-Taylor, Gailey, & Levinton, 1993; Mundt et al., 2007). Several studies found a decrease in formant frequencies in comparison with healthy individuals (Flint et al., 1993; France, Shiavi, Silverman, Silverman, & Wilkes, 2000; Mundt et al., 2007; Tolkmitt, Helfrich, Standke, & Scherer, 1982; Vicsi et al., 2012). This finding could be explained by psychomotor retardation that causes either tightening on vocal tract or lack of motor coordination (Cummins, Epps, & Ambikairajah, 2013; Flint et al., 1993; France et al., 2000; Quatieri & Malyska, 2012; Trevino, Quatieri, & Malyska, 2011).

Further acoustic speech measures, such as jitter, shimmer, are voice quality measures that are positively correlated with depression (Low et al., 2020; Ozdas, Shiavi, Silverman, Silverman, & Wilkes, 2004; Quatieri & Malyska, 2012; Vicsi et al., 2012). Indirectly-relevant features of voice properties (e.g., MFCCs or power spectral density) are also correlated with individuals' mood (Cummins et al., 2013; Cummins, Epps, Breakspear, & Goecke, 2011; France et al., 2000; Ozdas et al., 2004; Taguchi et al., 2018). Taguchi et al. (2018) investigated the differences in the MFCCs on individuals with and without depression and found evidence of higher levels of sensitivity and specificity of the second dimension of a MFCCs, concluding that this dimension could be a discriminatory factor between depressed and healthy patients and, consequently, a depression biomarker.

More consistent results were found with respect to another prosodic feature: speech rate. Cannizzaro et al. (2004) found a strong negative correlation between speech rate and a clinical subjective rating of depression. Studies using different sample sizes have concluded, in general, that speech rate is reduced in individuals with depression (Cannizzaro et al., 2004; Godfrey & Knight, 1984; Greden, Albala, Smokler, Gardner, & Carroll, 1981; Hardy, Jouvent, & Widlöcher, 1984; Mundt et al., 2007). Despite the value of speech rate as a potential biomarker of depression severity, it remains unclear whether the reduction in speech rate is an indicator of motor retardation or lower cognitive functioning (Alpert et al., 2001; Cannizzaro et al., 2004; Cummins, Scherer, et al., 2015; Trevino et al., 2011); additionally, speech rate could not present appropriate discriminatory evidence to be a single biomarker of depression (Cummins, Scherer, et al., 2015). Trevino et al. (2011) also concluded that an approach that includes pause and phone length information shows higher correlations than the global rate measurements for detecting depression symptoms.

Other suprasegmental speech measures were also found to have significant correlation with subjective measures of depression (Cannizzaro et al., 2004; Mundt et al., 2007). Total recording duration increased with depression severity due to more variable and longer pauses, which resulted in a decrease in speech to pause ratio (Alpert et al., 2001; Mundt et al., 2007). Percent of pause time is higher in the depressed group (Cannizzaro et al., 2004). Mundt et al. (2007, 2012) also revealed a significant decrease of the total recording duration, total pause time and number of pauses in patients that respond positively to depression treatment. By contrast, patients that do not respond to treatment presented smaller vocal acoustic changes or even no changes. As a result, the aforementioned measures could be considered as biomarkers to monitor treatment progress.

3.3 Articulatory studies of age effects on speech

Modifications in the supralaryngeal structures (particularly in the tongue) with aging, as well as decrease on motor tasks performance with increasing age (i.e, a generalized slowness, decrease coordination, a diminished performance level and on precise motor control), could considerably affect the speech production (Belmont, 2015; Goozée et al., 2005; Hermes et al., 2018). Further, Linville (2001) suggested that there is acoustic evidence of decreases in the extent of tongue movement across age, at least in males. However, direct, objective, physiological investigations of these changes are required (Goozée et al., 2005).

The articulatory studies concerning vowels' properties across lifespan are scarce and the majority of them focus on coarticulatory issues (Belmont, 2015; Zharkova, Hewlett, & Hardcastle, 2012). An articulatory study with 3D EMA for German suggested that the tongue body was especially affected by age, and the movements for the vowels were slower in the older speakers compared to the younger ones (Hermes et al., 2018). The results of a US study of anticipatory velar-vowel coarticulation and speech stability in English speakers who stutter and do not stutter across lifetime indicated an age effect, with progressive less coarticulation and an increase of speech stability with aging (Belmont, 2015). With US, Sonies, Baum, and Shawker (1984) studied the tongue movements, in healthy older and younger English speakers, during the repetition of the phonemes [a], [i] and [k], and observed a reduction in tongue retraction during the vowel [a] production for older speakers.

Regarding consonants, in a US study for Newfoundland English, De Decker and Mackenzie (2017) found a significant effect of age on the articulatory properties of /l/, with older speakers being more likely to exhibit distinctions in tongue gestures between the word-initial and word-final positions.

The literature suggests that the tongue may adapt naturally to age-related muscular changes by making subtle positional or movement changes within the oral cavity during speech production (Linville, 2001; Sonies, Baum, & Shawker, 1984), which may also reflect possible decline in oral motor precision as function of age (Linville, 2001).

To the best of our knowledge, for EP, there are no articulatory studies concerning age-related changes on vowel properties. The few articulatory studies available for EP are related with the production/configuration of nasal vowels through EMA or RTMRI (Cunha et al., 2020, 2019; Oliveira, Martins, et al., 2012; Oliveira et al., 2009).

3.3.1 Ultrasound imaging of speech

US imaging presents several advantages in comparison with other articulatory techniques (e.g., EPG, EMA, RTMRI): it is a non-invasive, safe, portable and fast technology that is commonly used to image the midsagittal surface contour of the tongue (Gonzalez, 2021; Lancia, Rausch, & Morris, 2015; Tabain, 2013), and it can contribute with important information for different areas of speech research (Akgul, Stone, & Maureen, 1999; Mozaffari, Wen, Wang, & Lee, 2019).

Despite US being a more affordable alternative for several contexts, enabling the acquisition of larger datasets, it demands adequate computational approaches for processing and analysis. US artifacts, corrupting noise, the presence of spurious edges, and the lack of a physiological reference are also challenges for the processing of US images (Akgul et al., 1999; Stone, 2005). Even though the tongue contours are visible, there are no hard structure references (i.e., US does not image internal articulators other than the tongue), making it difficult to determine an exact position of the tongue in the vocal tract (Tabain, 2013; Zharkova, Hewlett, & Hardcastle, 2011). The head and transducer holders help overcome these problems (Stone, 2005), but cannot be guaranteed to be re-fitted to the same location on a speaker's head in different trials (Scobbie, Lawson, Cowen, Cleland, & Wrench, 2011). For that, re-orienting images to a common coordinate system with a bite plane allows for some degree of normalization and it is more tolerant to error in the placing of the probe outwith the midsagittal plane (Scobbie et al., 2011).

Furthermore, often the US tongue contours do not provide information on the tongue tip and/or on the tongue root, due to the interfering presence of the mandible and hyoid bones (Gonzalez, 2021;

Tabain, 2013). Thus, the US image segmentation is strongly influenced by the quality of data (Noble et al., 2006). The speaker size and age also tend to influence the quality of the US tongue image (i.e., women and children tend to produce better images than men, younger subjects generally image better than older subjects, and thinner subjects tend to produce better images than larger subjects) (Tabain, 2013). Tabain (2013) suggests that these trends may be related to fat levels in the tongue and moisture levels in the mouth.

The lack of reference points and the anatomical differences also introduce a difficulty in comparing speech data across speakers, that is, comparing lingual articulation across age groups raises problems of normalization (Strycharczuk & Scobbie, 2017; Zharkova et al., 2011). There is no commonly accepted method for comparing tongue shapes among speakers (Barbier et al., 2015; Radisic, 2014).

Several research studies have been using US tongue images to investigate, on different languages, the articulatory correlates of vowel production (Comivi Alowonou et al., 2019; Ménard et al., 2013; Radisic, 2014; Strycharczuk & Scobbie, 2017). Most of them have focused on coarticulation aspects among children, adolescents and adults (Barbier et al., 2015; Belmont, 2015; Zharkova, Hewlett, & Hardcastle, 2008; Zharkova, Lickley, & Hardcastle, 2014). US studies on vowels could be divided in two types: qualitative (i.e., describe vowel articulation) and quantitative (i.e., measure some aspects/parameters of vowel articulation) (Radisic, 2014). Concerning quantitative methodologies, the articulatory parameters, tongue contour, height and advancement of the highest point of the tongue (i.e., TH and TA), lengths of posterior tongue surface (LPTS) and anterior oral cavity (LAOC) have been successfully used in US studies with vowels (Georgeton, Antolík, & Fougeron, 2016; S.-H. Lee, Yu, Hsieh, & Lee, 2015; Radisic, 2014; Song, 2017). Also ratios, such as Dorsum Excursion Index (DEI - quantifies the extent of tongue dorsum excursion towards the palate during an articulation) and Tongue Constraint Position Index (TCPI - represents the place of maximal constriction), have been used in coarticulation studies (Baghban, Zarifian, Adibi, Shati, & Derakhshandeh, 2020; Zharkova et al., 2014), which allow the comparison across speakers without the need to perform a normalization for vocal tract size (Zharkova, 2013). Nonetheless, to obtain these indices it is essential to ensure that the tongue curve in the image is visible between the mandible and hyoid bones shadows, and that both shadows are also visible (Zharkova, 2013) (as well as to obtain the LPTS and LAOC measures (S.-H. Lee et al., 2015)). Furthermore, the total vowel contour could also be analyzed through Smoothing Spline ANOVA, which allow to compare tongue contours and measure differences between segments (Davidson, 2006; De Decker & Nycz, 2012; Gonzalez, 2021; Mielke, 2015; Song, 2017; Turgeon, Trudeau-Fisette, Fitzpatrick, & Ménard, 2017).

US tongue imaging could be analyzed using static approaches, as mentioned above, but also through dynamic approaches. While the static approaches allow the comparison of tongue contours at specific articulatory landmarks (e.g., to study differences between vowels) (De Decker & Nycz, 2012; Mielke, Carignan, & Thomas, 2017), the dynamic approaches can be used to study the kinetics of speech sounds (i.e., allow measure articulatory data in an articulatory continuum) (Gonzalez, 2021). That is, dynamic approaches have used measures such as tongue displacement and velocities (Gonzalez, 2021; Strycharczuk & Scobbie, 2015). Nonetheless, data for dynamic approaches generally require more preparation than in static approaches (Gonzalez, 2021).

In summary, contrary to acoustics, where measuring formants to determine the vowel quality is a widely accepted approach, in speech studies with US various measurements have been suggested, but none have been standardized yet, since research with US is relatively recent (Radisic, 2014).

PART II

METHODS

This page intentionally contains only this sentence.

ACOUSTIC DATABASE OF EUROPEAN PORTUGUESE SPEECH

This chapter describes the research design of the acoustic database and includes details about the study region, participants, participants' tasks, corpus and data acquisition.

The objective of this work is to provide a first database containing all EP oral vowels produced in similar context (reading speech), and also semi-spontaneous speech (image description) collected from a large sample of adults in good health using standardized recording procedures. The collected data were used on speech acoustic studies over the lifespan.

4.1 Ethical issues

The study protocol of this cross-sectional study was submitted to and was approved by the Ethics Committee Centro Hospitalar São João/ Faculty of Medicine, University of Porto, Portugal (approval number N38/18) (see Appendix A), and all participants agreed and signed the informed consent term form before data collection (see Appendix B). There were no financial incentives provided for participation. The investigator recruited and collected data for all participants.

In order to guarantee the anonymity and confidentiality of the data, a numerical code was used to identify each participant. This code is common to the questionnaire and the recording session.

4.2 Study region

This study was carried out in the center of Portugal (namely the regions of Aveiro and Coimbra), due to the easy access to informants.

Taking into account the well-known Cintra (1971) classification of Portuguese dialects (*Nova proposta de classificação dos dialectos galego-portugueses*), the selected region is part of the central-southern Portuguese dialect (*português centro-meridional*), more specifically the centro-litoral dialect (*estremenho-beirões*), which covers *Estremadura* and part of *Beira Litoral* (see Figure 4.1) (Blayer, 2002; Segura, 2013).

The line that establishes the great division in Portuguese varieties between northern groups (i.e., northern Portuguese dialect (*português setentrional*)) and center-southern ones (namely the centro-litoral dialect) is the isophone that marks the southern limit of apico-alveolar sibilants (Cintra, 1971; Pérez, 2014; Segura, 2013) (line that crosses Portugal in the northwest-southeast direction, from the coastal region of Aveiro until the border with Spain, in Beira-Baixa - southern limit of the dark blue area in the Figure 4.1). In other words, the production of the phonemes /s/ and /z/ as apicoalveolar in the north and as predorsodental or dental in the south (Cintra, 1971; Ferreira, Carrilho, Lobo, Saramago, & Cruz, 1996; Segura, 2013). Additionally, the absence of phonological distinction between /b/ and

/v/, in favour of the first, which is produced, depending on the context, as occlusive [b] or fricative [β], is common in the northern region of Portugal (Cintra, 1971). However, the regions of Aveiro and Coimbra also present high percentages of merger of the bilabial and labiodental phonemes /b/ and /v/ (Matias, 1993; Pérez, 2014; I. A. Santos, 2003).

The central-southern Portuguese varieties mainly differ from each other in one segmental feature, namely the reduction of the diphthong [ej] to [e] in the interior center and south (yellow area); while the litoral-center (brown area) maintains the diphthong [ej], produced as [ɐj] (as in Lisbon) (Cintra, 1971; Segura, 2013; Vigário & Frota, 2003). This feature was also confirmed for the region under study by Pérez (2014). Nevertheless, the boundaries between the two dialect regions are not evident.

More recently, Segura (2013) suggested a new syntactic phonetic feature to add to the features of Cintra (1971), which also divides the country into two areas: the insertion of a semivowel between two central vowels, belonging to different words, in order to break the hiatus (e.g., "**a á**gua" [ɐj'aɣwɐ]). This is common in the north and center of Portugal, and Matias (1993), in a sociolinguistic study, also reported that speakers of Aveiro city produced an [j] anti-hiatus in 75% of the linguistic productions.

Some authors point that the central region of Portugal (including Beiras) is a transitional area between the north (i.e., a conservative area) and south (Blayer, 2002; Boléo & Silva, 1961; Brissos, 2020; Cuesta & Luz, 1980). Additionally, most studies focus on these extremes (Brissos, 2020), as they consider *Beira Litoral* as a dialectal region without distinctive dialect/phonetic features (Boléo & Silva, 1961; Salema, 2007).

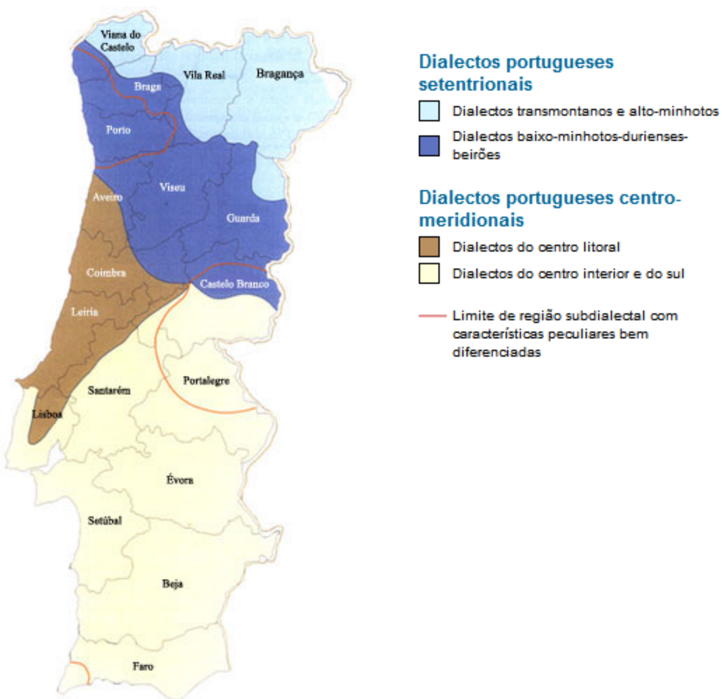


Figure 4.1: Classification of Galician-Portuguese dialects, by Cintra (1971) (adapted from Segura & Saramago (2001)).

4.3 Participants selection

A convenience sampling strategy was used to recruit healthy adults over the age 35. To ensure an equitable distribution of participants, the following age groups were covered: [35-49], [50-64], [65-79] and ≥ 80, with at least ten females and ten males in each group.

The purpose of this age grouping was to allow a comparison between individuals at different stages in the aging process. The main assumptions for this age division rely on the chronological definition of older people, and also on the reported patterns of physiological changes along the aging continuum (cf. Section 2.1.1) (He et al., 2016; Mautner, 2011; World Health Organization, 2007) . The youngest group [35-49] represents adults prior to the onset of physical aging that involve changes of the vocal fold tissue (Mautner, 2011). The age groups [50-64] and [65-79] were defined based on the World Health Organization definition of older persons (He et al., 2016). The oldest group ≥ 80 represents healthy adults who are at an age where changes to the vocal fold tissues tend to be more prevalent in the literature (Mautner, 2011).

In addition to the age grouping, gender was also considered due to the substantial gender differences in the extent and timing of the aging process (Linville, 2001; Makiyama & Hirano, 2017; Mautner, 2011; Pontes et al., 2005; Schötz, 2006).

Participants were recruited through personal contacts and through snowball technique in the community, and in Senior Universities (Universidade Sénior de Oliveira do Bairro, Universidade Sénior de Cacia, Academia dos Saberes de Aveiro, Universidade Sénior Gafanha da Nazaré - Ílhavo, Universidade Sénior de Águeda), day centers and nursing homes (Centro Social de Oiã and Santa Casa da Misericórdia do Concelho de Oliveira do Bairro) from the central region of Portugal, and also in the University of Aveiro.

Participants were included if they: were native Portuguese speakers (who lived in a non Portuguese-speaking country for no more than 5 years); lived in Aveiro or Coimbra region; had no history of speech-language impairment; had no severe hearing problems; and no history of neurological disorders or head/ neck cancer. Also, the participant inclusion criteria consisted of non-hospitalized and free of upper respiratory tract infection for 3 weeks prior to the experiment. Participants were excluded if they: 1) were current smokers or had smoked within the previous 5 years; 2) self-reported poor general health; 3) wore hearing aids, 4) had received speech and language therapy, and 5) reported that their voice was different than usual on the day of testing (eg, having a cold or allergy symptoms). Participants who exhibited any observable sign of speech, voice, or severe hearing problems as assessed by a speech pathologist on the moment of recording and those who were unable to follow directions were excluded.

The exclusion criteria were defined considering factors that affect vocal quality (Guimarães, 2007; Sataloff et al., 2017), as well as the criteria considered in studies of a similar nature (Dehqan et al., 2013; Goy et al., 2013; Guimarães & Abberton, 2005; Mautner, 2011; Sebastian et al., 2012; Selent, 2014). 48 participants were excluded as they presented one or more exclusion factors. A total of 144 participants were included in this database. Figure 4.2 shows the distribution of the participants by age group and gender.

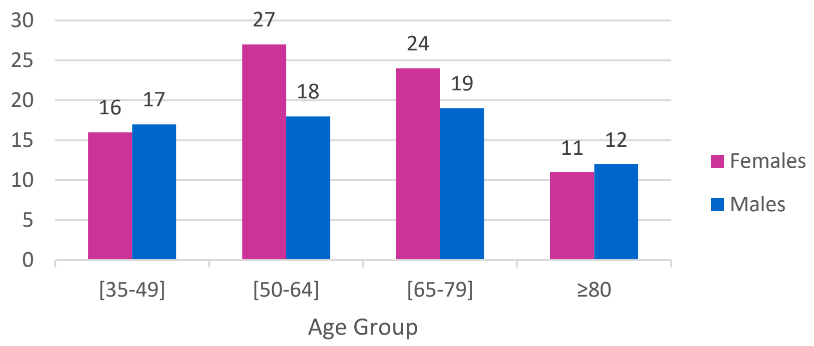


Figure 4.2: Number of participants recorded by gender and age group.

4.4 Instruments

For the purposes of this database a background questionnaire was developed, which intends to collect information concerning age, gender, educational level and habits (see Appendix C).

The background questionnaire was designed considering the inclusion and exclusion criteria previously defined. A pre-test was carried out with 10 participants (5 females and 5 males aged between 28 and 86). The participants answered the questionnaire and then its duration, ambiguous or difficult questions were analyzed, as well as if the questions allowed to obtain the intended information. Considering the pre-test results, some questions were reformulated in the final version of the questionnaire (see Appendix C).

The background questionnaire contained socio-demographic items (gender, academic qualifications, professional status, place of birth and city where he/she resides), medical and voice related history (self-evaluation of health status, reproductive phase, previous diagnoses of clinical conditions that may affect vocal quality, professional use of voice), support device needs (hearing aids, glasses or contact lenses, dental prostheses or orthodontic appliances) and surrounding/environmental conditions and habits (exposure to adverse conditions, smoking habits, alcohol, water and caffeine consumption).

The participants also fulfilled a self-reported instrument concerning anxiety and depressive symptomatology ("Hospital Anxiety and Depression Scale" (HADS) (Zigmond & Snaith, 1983), Portuguese version of Pais-Ribeiro et al. (2007)) (see Appendix D). The instrument HADS was selected considering the nature of the phenomena intended to study, as well as their psychometric properties. Furthermore, the authors of the Portuguese version of the HADS authorized its use in this study.

The HADS was used to evaluate anxiety and depression symptoms. In the context of the HADS, anxiety refers to generalized anxiety and depression refers to a loss of pleasure (Dietrich, Abbott, Gartner-Schmidt, & Rosen, 2008). HADS is not a time consuming instrument, and has been largely used in research studies and in clinical settings with non-psychiatric populations (Pais-Ribeiro et al., 2007). It presents good internal consistency, sensitivity and specificity and concurrent validity (Bjelland, Dahl, Haug, & Neckelmann, 2002). HADS is divided into an Anxiety subscale (HADS-A) and a Depression subscale (HADS-D) with seven items each that address symptoms from the preceding week. Each item has a 4-point Likert score scale with a minimum value of 0 and a maximum value of 3. Higher scores represent higher levels of anxiety and depressive symptoms. The HADS manual provides cut-off scores indicating mild (8–10), moderate (11–14), or severe (15–21) anxiety or depression (Dietrich et al., 2008; Pais-Ribeiro et al., 2007; Zigmond & Snaith, 1983). Following the cut-offs,a score of 0 – 7 for each subscale could be regarded as being without anxiety or depression symptoms (Pais-Ribeiro et al., 2007). So, 7 is the maximum value for the normal range.

The Portuguese version of HADS, which was validated with a sample of 1331 participants, is a reliable and valid instrument for assessing anxiety and depression and present metric properties similar to the general studies in other languages. The internal consistency of HADS-A was $\alpha = 0.76$, and for HADS-D was 0.81 (Pais-Ribeiro et al., 2007).

4.5 Corpus

The corpus consists of two types of data: the first refers to segmental and the second to suprasegmental data.

4.5.1 Segmental corpus

The speech corpus consisted of 36 words, with the EP oral vowels [i], [e], [ɛ], [a], [o], [ɔ], [u] in stressed position and the vowels [ɨ] and [ɐ] in unstressed position. Each vowel was produced in a disyllabic sequence, mostly C**V**.CV (C-consonant, V-vowel)[1] (e.g. "pato", *duck*), where C was a voiced/ voiceless stop consonant ([p], [t], [k], [b], [d], [g]) or a voiced/ voiceless fricative consonant ([f], [s], [ʃ], [v], [z], [ʒ]). To facilitate the vowel segmentation, words with voiceless consonants are preferred (Escudero et al., 2009). However, due to the difficulty to find real words, it was also necessary to include words with voiced consonants. The list of 36 words used in this study is presented in Table 4.1 (in International Phonetic Alphabet (IPA)), and contains four words per vowel.

Table 4.1: List of words per vowel (International Phonetic Alphabet). The symbol * indicates that the pictogram of this word presented a naming percentage of accuracy lower than 70%. (adapted from Albuquerque et al. (2020))

Oral vowels		**Words**			
stressed	**[i]**	[ˈfitɐ] (*ribbon*)	[ˈbiku] (*beak*)	[ˈfigu] (*fig*)	[ˈpizɐ]* (*pizza*)
	[e]	[ˈseʃtɐ] (*basket*)	[ˈdedu] (*finger*)	[ˈpezu] (*weight*)	[ˈzebrɐ] (*zebra*)
	[ɛ]	[ˈsɛtɨ] (*seven*)	[ˈtɛtu] (*ceiling*)	[ˈsɛtɐ] (*arrow*)	[ˈʃɛkɨ] (*check*)
	[a]	[ˈʃavɨ] (*key*)	[ˈfakɐ] (*knife*)	[ˈgatu] (*cat*)	[ˈpatu] (*duck*)
	[ɔ]	[ˈkɔpu] (*glass*)	[ˈbɔtɐ] (*boot*)	[ˈfɔkɐ] (*seal*)	[ˈtɔʃɐ]* (*torch*)
	[o]	[ˈbokɐ] (*mouth*)	[ˈkoku] (*coconut*)	[ˈposu] (*well*)	[ˈgotɐʃ] (*drops*)
	[u]	[ˈʃuvɐ] (*rain*)	[ˈʃupɐ]* (*lollipop*)	[ˈkubu]* (*cube*)	[ˈʒubɐ]* (*mane*)
unstressed	**[ɐ]**	[kɐˈfɛ] (*coffee*)	[ʃɐˈpɛw] (*hat*)	[pɐˈtĩʃ] (*rollerblades*)	[pɐˈpɛɫ] (*paper*)
	[ɨ]	[bɨˈbeɾ] (*to drink*)	[dɨˈdaɫ] (*thimble*)	[pɨˈdaɫ] (*pedal*)	[pɨʃˈkaɾ]* (*to fish*)

The speech stimuli were carefully chosen to allow easy and accurate formant measures since the vowel context is restricted to stop and fricative consonants. The corpus was also designed to collect data over the lifespan. Thus, the words were therefore chosen to be familiar to all ages, and also,

[1]To clarify, the first **V** correspond to the vowel target.

easily represented by images (to avoid the interference of reading abilities in the production of words) (Eichhorn et al., 2018).

A pilot naming study was conducted for the selection of these images. 63 pictures were selected from color pictograms of ARASAC (Palao, 2017) and presented to a group of 10 participants (5 males and 5 females), ranging from ages 28 to 86. The results indicated that adult participants were able to properly name most of the pictures (percentage of accuracy equal or higher than 70%, except for 6 images) (see Table 4.1). However, these stimuli remained in the study due to the difficulty in finding alternative words that met the previously defined criteria. The stimuli were embedded in a carrier sentence "Diga ... por favor" (*"Say ... please"*).

4.5.2 Suprasegmental corpus

Semi-spontaneous speech samples were collected from the participants using a standard Boston Diagnostic Aphasia Examination (Goodglass & Kaplan, 1983) picture description task, with the standardized picture "Cookie Theft" stimulus shown in Figure 4.3. The participants were instructed to describe the picture (Morgan & Rastatter, 1986; Pakhomov et al., 2011).

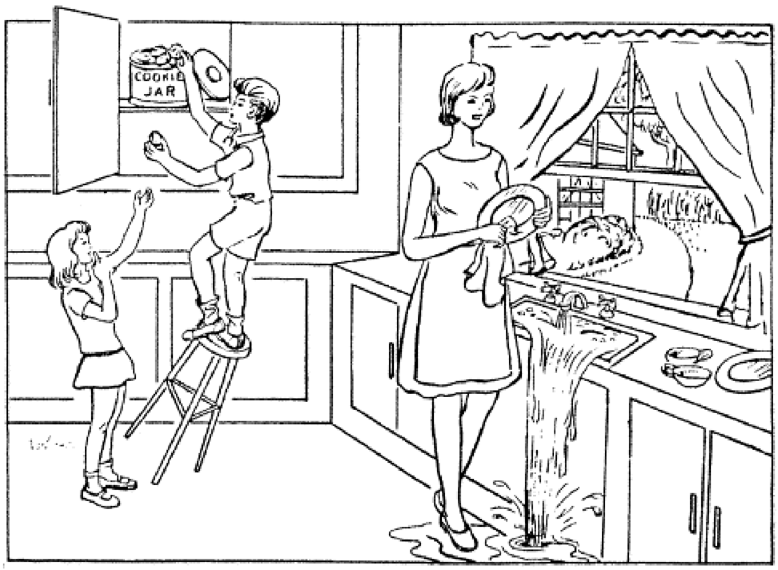


Figure 4.3: The "Cookie Theft" picture from the Boston Diagnostic Aphasia Examination, in Goodglass and Kaplan (1983).

4.6 Data acquisition

The data acquisition was conducted between February 28, 2018 and January 30, 2019. The same researcher was present in all data acquisition sessions. The participants filled in the questionnaires (i.e., background questionnaire and the HADS), which took approximately 15 minutes to complete. In exceptional cases, due to low literacy or visual difficulties, the questions were presented orally by the researcher.

Regarding the corpus acquisition, all recordings took place in quiet rooms using an AKG C535 EB cardioid condenser microphone connected to a USB external 16-bit sound system (PreSonus AudioBoxTM USB), with a sampling rate of 44100 Hz. The participants were seated at a table and the microphone was adjusted to each participant and positioned at an approximately 15-20 cm distance from the mouth.

The sentences were randomized and presented individually on the computer screen using the software system SpeechRecorder (Draxler & Jänsch, 2004, 2017), where picture and orthographic words could be viewed simultaneously (Figure 4.4).

Figure 4.4: Visualization of target word picture (pictogram of ARASAC (Palao, 2017)) and the carrier sentence in the SpeechRecorder program (Draxler & Jänsch, 2004, 2017).

Three practice words in the same carrier sentence were used at the beginning to adjust recording levels. Thus, participants were asked to read the sentences at comfortable pitch and loudness level, after familiarizing themselves with the structure of the sentences. During recording, stimuli that were mispronounced or occurred with intense background noise were repeated. Additionally, they could take a break at any time they wished and each speaker attended a single recording session.

Each carrier sentence was repeated three times. Thus, each participant produced 12 repetitions of each vowel, in a total of 108 productions by speaker, and needed approximately 15 minutes to complete this task.

The participants were also instructed to describe the "Cookie Theft" picture (Goodglass & Kaplan, 1983) at comfortable pitch and loudness level, after familiarizing themselves with the image, in order to obtain induced spontaneous speech (Morgan & Rastatter, 1986; Pakhomov et al., 2011). The image was also presented on the computer screen with software system SpeechRecorder (Draxler & Jänsch, 2004, 2017). The instruction given to participants was as follows: "Tell me everything you see in this picture.". This task took approximately 30 seconds to complete.

This page intentionally contains only this sentence.

SEGMENTATION OF THE ACOUSTIC DATA SET

This chapter includes information about the extraction of the acoustic measures at both segmental and suprasegmental levels.

The basic steps before segmentation were the following: speech recordings were uploaded to a computer, and saved into separate files (by word and image description) by the software system SpeechRecorder (Draxler & Jänsch, 2004, 2017). The files of each participant were saved in a folder identified with its numerical code.

For a long time the processing of data from sizeable speech corpora was limited by technical shortcomings, which made the data segmentation difficult and time-consuming, but automated systems nowadays allow to perform these tasks easier and in a less amount of time (Elvira-García, 2014; Van der Harst & Van de Velde, 2014; C. I. Watson & Evans, 2016). Despite the availability and use of programmable tools, it was only possible to annotate part of the 144 valid participants in the database, as manually correction of the segmentation is still a time-consuming task (C. I. Watson & Evans, 2016). The selection of the participants for manual correction of the segmentation was based on the sound quality of the recordings and tried to ensure an equitable distribution of participants by the age groups previously defined (i.e., [35-49], [50-64], [65-79], and ≥ 80).

In addition, the sociodemographic characterization of the speakers who are part of the cured databases is presented at the end of this chapter.

5.1 Segmental annotation

The recorded data was first automatically segmented at word and phoneme level using WebMAUS General for Portuguese language (PT) (Kisler, Reichel, & Schiel, 2017; Schiel, 1999) and then imported into Praat speech analysis software (Version 6.1.08) (Boersma & Weenink, 2012), so that four trained analyzers could manually check and correct, when necessary, the accuracy of the vowel boundaries. Thus, the segmentation and labeling of each carrier word, vowel target and flanking consonants were manually verified, over the digitized sound wave (see Figure 5.1). The start and end points of the vowel were determined by finding the first and last periods that had considerable amplitude and whose shape resembled that of more central periods, with both points of the selection chosen to be at a positive zero crossing of the waveform.

The analyzers revised 12 204 recordings of 113 speakers, as each participant produced 12 repetitions of each vowel, in a total of 108 productions by speaker (113 participants x 36 words x 3 repetitions = 12 204 recordings). Note that, the 108 productions per speaker are the result of 36 words repeated 3 times.

A total of 736 recordings were discarded (approximately 6% of trials) due to problems with the recordings (e.g., clipping, noise, misread, hoarseness or vocal fry) or vowel reduction (vowel [ɨ] was the most deleted vowel, mostly in the context of “pescar” ([pɨʃˈkaɾ] - *to fish*)). In the present study the vowel [ɨ] deletion was observed in 26.7% (359 deletions) of the vowel occurrences. Previous studies of EP spontaneous speech have reported higher occurrences of [ɨ] deletion, around 40% to 70% (Rodrigues, 2016). Studies on other languages have also indicated higher levels of vowel deletion in spontaneous speech comparing with reading context (Adda-Decker, Boula De Mareüil, & Lamel, 1999; Munson, 2007).

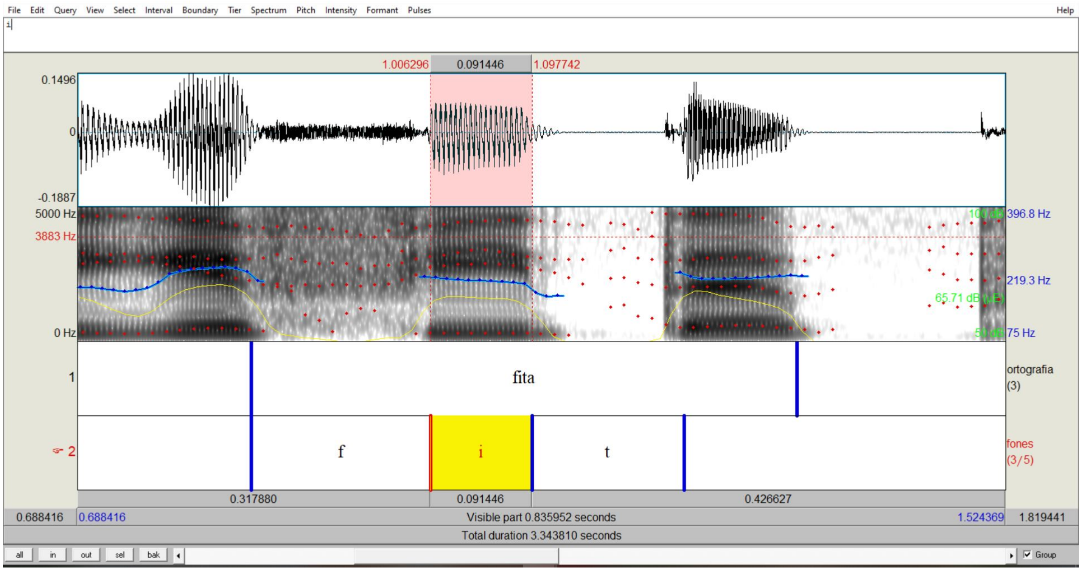


Figure 5.1: Annotation of vowel target, flanking consonants and carrier word in Praat software. An example of the word "*fita*" - [ˈfitɐ].

5.2 Suprasegmental annotation

Concerning semi-spontaneous speech, the speech and pauses were labeled, and in the speech intervals, the vowel onsets were also detected, through automatic Praat scripts.

The Praat Script Syllable Nuclei v2 (de Jong & Wempe, 2009)[1] was used to automatically detect silent pauses of over length 250 ms (Cannizzaro et al., 2004) and create textgrid files. The automated alignments of silent pauses were manually checked by two trained analyzers, who verified the accuracy of pause and speech intervals, and also labelled intervals with speaker or environmental noise. The intervals were labeled as: pause (breathing sound was considered as silent pause), speech, verbal non lexical (i.e., filled pauses), noise (i.e., noise that occurs during the speaker's pauses), vocal non lexical (i.e., laughter, coughing or other human noises), and speech with noise (i.e., speech intervals with environmental noise that could affect the acoustic measurements) (Pellegrino, 2019; Schuller et al., 2013)

Vocal non lexical phenomena were considered as pause time, while verbal non lexical phenomena were not included in the present analysis (Pellegrino, 2019; Tuomainen, Hazan, & Taschenberger, 2019). Speech intervals with noise were not counted for further analysis, and also the beginning and

[1]The Praat Script Syllable Nuclei v2 was modified by Hugo Quené, Ingrid Persoon, & Nivja de Jong and is available in: https://sites.google.com/site/speechrate/Home/praat-script-syllable-nuclei-v2.

ending of all recordings were not pondered in the analysis due to sentence initial and final acoustic variability (a total of 7% of the speech intervals were excluded).

Regarding the syllables spoken, an adapted Praat script of the BeatExtractor (Barbosa, 2006, 2010) was used to detect vowel onset using a beat wave (a normalized and band-specific amplitude). The cut-off frequency was defined automatically. The thresholds were 0.1 and 0.06, the filter was defined as Butterworth and the technique was Amplitude. The total number of syllables were automatically obtained through the sum of all vowel onset detected within all valid speech intervals per speaker. A random check was done to verify the vowel onsets and confirm the script performance.

5.3 Acoustic measurements

A set of acoustic measurements were extracted from the recording data. The following procedures were adopted in the extraction of data for the segmental and suprasegmental domains.

5.3.1 Segmental parameters

Acoustic parameters (F0 and formant frequencies) were automatically extracted from the central 40% of each target vowel using Praat scripts (Escudero et al., 2009). Median F0 value of the vowels was estimated with the cross-correlation algorithm, which is especially suitable for measuring short vowels (Escudero et al., 2009). Median F0 value was taken from the central 40% of each target vowel, which minimizes the impact of flanking consonants on the F0; in addition, taking the median F0 values rather than the mean, reduces the effect of F0 measurement errors (Escudero et al., 2009). The pitch range for the analysis was set to 60 - 400 Hz for men and 120 - 400 Hz for women. If the analysis failed on any of the speaker's vowel tokens, that token was excluded (31 vowels, most of them produced by an 80-year-old woman, which was excluded from all the analysis).

Burg linear predictive coding (LPC) algorithm, as provided by Praat, was used to compile values for F1, F2 and F3. A procedure (adapted from Escudero et al. (2009) and previously used in Oliveira, Cunha, et al. (2012) and Albuquerque et al. (2014)) was applied to optimize the formant ceiling for a certain vowel of a certain speaker. The first three formants were determined 201 times for each vowel, for all ceilings between 4500 and 6500 Hz for female and between 4000 and 6000 Hz for male, in steps of 10 Hz. The chosen ceiling was the one that yielded the lowest variation. Thus, for each vowel produced by each speaker there is only one "optimal ceiling" (for more details see Escudero et al. (2009)).

The duration measurements were computed from the label files with reference to the beginning and the ending points of each vowel. Vowels with duration values shorter than 20 ms were excluded (8 vowels), and outliers that exceeded 2.5 standard deviations from the mean for particular speaker by F0 and from their gender x vowel mean by formant frequency were also excluded from this analysis (Eichhorn et al., 2018; Smiljanic & Gilbert, 2017). In this study, the measurements for duration, F0, F1-F3 were manually checked for possible extraction errors and these procedures yielded in the removal of 695 outliers (nearly 1.5% of the total data).

Reliability in vowel segmentation

To determine inter- and intra-rater reliability of the measures, 36 textgrids of each analyzer (1 textgrid randomly selected from each word) were relabelled by all analyzers. Thus, 144 (1.2%) of a total of 12 204 textgrids (113 participants * 36 stimulus * 3 repetitions) were manually relabelled for reliability by the four judges. The scripts to obtain vowel duration and formant frequencies were then readministered.

Inter and intra-rater reliability was assessed using the intraclass correlation coefficient (ICC) and the two-way mixed model (the raters were considered fixed) with an absolute agreement definition.

Reliability among the raters was considered excellent, with ICC values >0.952 for all vowels/ acoustic parameters (duration, F0, F1, F2, F3), except F1 of [ɨ] where ICC was 0.846, but still considered good reliability (Shrout & Fleiss, 1979).

To assess intra-rater reliability, a random sample of 36 textgrids (one of each stimuli) was manually rechecked by the same rater. Again, reliability was excellent with ICC values >0.909 for all vowels/ acoustic parameters (Shrout & Fleiss, 1979).

5.3.2 Suprasegmental parameters

Acoustic parameters (speaking F0 and HNR) were automatically extracted from the valid speech intervals, using the Praat script ProsodyDescriptor (Barbosa, 2013), with the F0 threshold of 75 - 400 Hz for males, and 120 - 600 Hz for females. The script extracted and calculated the following parameters: speaking F0 mean (semitones (re 1 Hz)) and HNR. Each value was considered to obtain the average of speaking F0 for each participant. The F0 scale with values in semitones relative to 1 Hz was converted to Hertz.

Additionally, based in the suprasegmental annotation of speech, pauses and syllables, several rhythm measures were computed, such as speech duration, pause duration, percent pause time, speech pause ratio, number of pauses, speech rate and articulatory rate.

5.4 Sample characterization

The speech data (vowel tokens and spontaneous speech) of 112 participants were processed, and this speech database was analyzed in future acoustic studies (see Chapters 7, 8, 9 and 10).

A total of 112 native Portuguese speakers (56 men and 56 women), from the central region of Portugal, aged between 35 and 97 (females: 61.6 ± 15.8 years; males: 62.6 ± 15.5), were included in this speech database. Despite the exclusion of an 80-year-old woman, the sample is almost balanced in terms of age and gender: [35-49] (15 men, 15 women), [50-64] (15 men, 15 women), [65-79] (15 men, 16 women), and ≥ 80 (11 men, 10 women). All participants self-reported to live independently and actively interacting with family and friends within the community, except 2 males and 1 female in the age group ≥ 80, who lived in a nursing home. Furthermore, 1 male and 1 female attended the day center.

Table 5.1 shows the demographic information for the female and male participants for the four age groups. The majority of the participants have more than 12 years of formal school (37.5%), followed by participants with 1 to 4 years of school (32.1%), and participants with 5 to 12 years of school (30.4%). As the age increases, the number of participants with higher educational level decreases. Also related with this issue, 26.8% of the speakers needed support of the researcher to answer the questionnaires, mostly in the last age groups.

Table 5.1: Demographic and lifestyle characteristics of the speakers by age group and gender.

Characteristics	Males				Females			
	[35-49] N=15	**[50-64]** N=15	**[65-79]** N=15	**>=80** N=11	**[35-49]** N=15	**[50-64]** N=15	**[65-79]** N=16	**>=80** N=10
Mean age (age range) (years)	44.3 (35-49)	56.3 (60-64)	70.7 (65-79)	85.2 (81-97)	41.1 (35-49)	57.5 (50-63)	70.4 (67-77)	84.1 (80-90)
Years of school								
1 to 4 (%)	6.7	13.3	26.7	100.0	6.7	26.7	26.7	70.0
5 to 12 (%)	40.0	40.0	40.0	0.0	26.7	33.3	40.0	20.0
>12 (%)	53.3	46.7	33.3	0.0	66.7	40.0	33.3	10.0
City (%)								
Águeda (%)	0.0	0.0	13.3	0.0	0.0	13.3	18.8	20.0
Aveiro (%)	13.3	13.3	40.0	0.0	33.3	13.3	37.5	20.0
Ílhavo (%)	6.7	0.0	6.7	0.0	0.0	6.7	12.5	0.0
Oliveira do Bairro (%)	60.0	73.3	40.0	100.0	53.3	66.7	25.0	60.0
Others * (%)	20.0	13.4	0.0	0.0	13.4	0.0	6.3	0.0
Live outside of Portugal								
No (%)	66.7	60.0	46.7	54.5	86.7	100.0	75.0	70.0
Portuguese-speaking country (%)	13.3	20.0	40.0	36.4	6.7	0.0	6.3	30.0
Non Portuguese-speaking country**(%)	20.0	20.0	13.3	9.1	6.7	0.0	18.8	0.0
Reply to questionnaires								
by oneself (%)	100.0	100.0	80.0	0.0	93.3	80.0	68.8	30.0
with the support of the researcher (%)	0.0	0.0	20.0	100.0	6.7	20.0	31.3	70.0
Self-reported general health								
Average (%)	20.0	40.0	53.3	81.8	13.3	60.0	37.5	60.0
Good (%)	60.0	53.3	46.7	18.2	73.3	40.0	62.5	40.0
Excellent (%)	20.0	6.7	0.0	0.0	13.3	0.0	0.0	0.0
Emotional problems, yes (%)	26.7	13.3	20.0	18.2	33.3	40.0	56.3	10.0
Chronic respiratory disorder, yes (%)	6.7	0.0	0.0	9.1	20.0	20.0	18.8	10.0
Heartburn or reflux, yes (%)	6.7	0.0	20.0	0.0	6.7	6.7	31.3	20.0
Mild hearing impairment, yes (%)	6.7	0.0	13.3	18.2	0.0	0.0	0.0	40.0
Dental prostheses, no (%)	93.3	93.3	40.0	12.2	80.0	66.7	6.3	20.0
Daily medication, yes (%)	33.3	40.0	93.3	100.0	26.7	80.0	100.0	100.0
Menopause, yes (%)	–	–	–	–	13.3	93.3	100.0	100.0
Former smoker*, yes (%)**	6.7	13.3	40.0	0.0	6.7	6.7	12.5	0.0
Regular alcohol consumption, no (%)	46.7	53.3	33.3	45.5	100.0	86.7	75.0	80.0
Regular caffeine consumption, no (%)	20.0	13.3	20.0	72.7	40.0	13.3	43.8	40.0
Daily water consumption								
no (%)	6.7	0.0	6.6	9.1	6.6	0.0	6.3	0.0
Less than 1l (%)	40.0	60.0	46.7	18.2	66.7	40.0	56.2	70.0
1l or more (%)	53.3	40.0	46.7	72.7	26.7	60.0	37.5	30.0
Vocal training, yes (%)	13.3	13.3	46.7	0.0	6.7	0.0	18.8	20.0

*Albergaria a velha, Anadia, Condeixa-a-nova, Ovar, Santa Maria da Feira and Vagos
**1-5 years
*** >5 years ago

The speech database included participants who reported concussions in the past with no lasting consequences (1.8%), participants who reported having heartburn or gastrointestinal reflux (11.6%), or a chronic respiratory disorder (10.7%) or endocrine disorders (1.8%). And still, participants who reported (having or having had) emotional alterations such as depression (28.6%) or who have had any thorax surgery, head or neck surgery, oral or facial surgery (9.8%). 11.6% of the participants are ex-smokers and 15.2% referred that they have some type of vocal training (e.g., singing lessons).

54.5% of the speakers did not use any type of dental prostheses, beyond these 40.2% used fixed or removable dental prostheses and 5.3% of the speakers reported having few or no teeth.

Concerning the reproductive phase, females who reported hysterectomy were considered to be in the menopause, which is the case of one female in the first age group and another in the second. Thus, the majority of the female participants in the age group [35-49] were in the reproductive phase (86.7%), while in the age group [50-64] 93.3% were in the menopause.

Regarding alcohol consumption, 65.2% of the speakers reported no consumption or only consuming occasionally (i.e., no daily or regular alcohol consumption). The other participants (34,8%) reported drinking alcohol once to twice a day. 31.3% of the speakers do not consume caffeine on a daily basis, while 59.8% of the speakers consume 1 to 2 coffees per day, and 8.9% drink more than two per day.

ARTICULATORY DATABASE OF EUROPEAN PORTUGUESE SPEECH - ULTRASOUNDS

This chapter describes the research design of the articulatory database and includes details about the participants, background questionnaire, corpus and data acquisition. This work provides a database containing all EP oral vowels produced in similar contexts (pseudowords and isolated) collected from young and old adults through standardized recording procedures of US tongue imaging synchronized with audio. This database allowed to examine the speech articulatory changes with aging in EP.

6.1 Ethical issues

The study protocol of this cross-sectional study complies with the provisions of the General Data Protection Regulation (GDPR), and it was approved by the Ethics Committee of Escola Superior de Enfermagem de Coimbra, Portugal (approval number 639/12-2019) (see Appendix E). All participants agreed and signed the consent form before participating in the study (see Appendix F). There were no financial incentives provided for participation.

In order to guarantee the anonymity and confidentiality of the data, a numerical code was used to identify each participant. This code is common to the questionnaire and the recording session (US data and audio data).

6.2 Participants

A convenience sampling strategy was used to recruit young (between the ages 18 to 35) and old adults (over the age 55), to test ages with more expected distinctive characteristics.

Participants were recruited through personal contacts and through snowball technique in the community, and in the University of Aveiro. All of them were in good health and with no reported history of neurological disorders or diseases, or any speech, language or hearing difficulties. All participants were native Portuguese speakers, and reported no previous history of speech-language impairments, head/ neck cancer and/or neurological disorders. They were free of upper respiratory tract infection, and were excluded: 1) if they were current smokers or had smoked within the previous 5 years; 2) if they reported poor general health; and 3) if they wore hearing aids. As in the acoustic database, the exclusion criteria were defined considering factors that affect vocal quality (Guimarães, 2007; Sataloff et al., 2017), as well as the criteria considered in studies of a similar nature (Belmont, 2015; Carl, 2018).

One old female speaker was excluded of the database because she was a smoker. Therefore, US data were collected from a convenience sample of 31 EP healthy speakers (11 men and 20 women), ages ranged from 21 to 73, 15 young (6 males; 9 females) and 16 old adults (5 males; 11 females). All speakers lived in the region of Aveiro, except one older female who lived in Viseu. Data collection took place in a COVID-19 pandemic year, which impacted the recruitment of a larger number of participants, particularly older people.

6.3 Background questionnaire

For the purposes of this database a background questionnaire was developed, which also took into account the inclusion and exclusion criteria previously defined. The background questionnaire of the US study (see Appendix G) contains socio-demographic items (gender, academic qualifications, place of birth and city where he/she resides), anatomical characteristics (i.e., weight and height), medical and voice related history (self-evaluation of health status, previous diagnoses of clinical conditions that may affect vocal quality or swallow, support device needs (hearing aids, glasses or contact lenses, dental prostheses or orthodontic appliances), characteristics of the dentition and smoking habits.

6.4 Corpus

The corpus consisted of all EP oral vowels [i], [e], [ɛ], [a], [o], [ɔ], [u], [ɨ] and [ɐ] in pseudoword context and in isolated context. The pseudoword list contained ˈpV.Cv sequences (started with the labial voiceless stop consonant [p]), where C was balanced for the place of articulation using the voiceless stop consonants [p], [t] and [k], and V was all EP oral vowels in stressed position. The last vowel (i.e., v) corresponds only to the vowels [u], [ɨ] and [ɐ]. The first consonant was limited to non-lingual consonant, in order to avoid lingual coarticulatory effects during the vowel production (Carl, 2018). In contrast to the acoustic database, as this speech corpus consisted in pseudowords, it was possible to limit the context to voiceless consonants, in order to simplify the vowel target boundaries identification (Escudero et al., 2009).

The stimuli were embedded in a carrier sentence “Em pVCv temos V” (*In pVCv we have V*), where the last V was considered an isolated vowel and corresponded to the same vowel that occurred in the stressed position of the pseudoword (e.g. “Em pêca temos ê”). For each vowel, three different pseudowords were selected. The list of 27 pseudowords used in this study is listed in Table 6.1 in IPA symbols, however the pseudowords were presented to the speakers in the respective writing convention of EP (e.g., [ˈpekɐ] - “pêca”).

6.5 Data acquisition

The data acquisition was conducted in the Institute of Electronics and Informatics Engineering of Aveiro (IEETA), between July 29, 2021 and September 18, 2021. As one of the major challenges of the US image refers to the processing, visualization and analysis of the data collected, an appropriate approach to these aspects was needed. Thus, before the data acquisition that devises the current articulatory database, pilot experiments were carried out in order to assess the data acquisition protocol and contribute to the development of the US image segmentation method used (i.e., automatic extraction

Table 6.1: List of pseudowords per vowel (International Phonetic Alphabet). (adapted from Albuquerque, Valente, et al. (2022))

Vowels		Pseudowords		
front	**[i]**	[ˈpipɨ]	[ˈpitɐ]	[ˈpikɐ]
	[e]	[ˈpepɨ]	[ˈpetu]	[ˈpekɐ]
	[ɛ]	[ˈpɛpɨ]	[ˈpɛtɐ]	[ˈpɛku]
central	**[ɨ]**	[ˈpɨpɨ]	[ˈpɨtɐ]	[ˈpɨkɐ]
	[ɐ]	[ˈpɐpɨ]	[ˈpɐtɐ]	[ˈpɐkɐ]
	[a]	[ˈpapɨ]	[ˈpatɐ]	[ˈpaku]
back	**[u]**	[ˈpupɨ]	[ˈputu]	[ˈpukɐ]
	[o]	[ˈpopɨ]	[ˈpotu]	[ˈpoku]
	[ɔ]	[ˈpɔpɨ]	[ˈpɔtu]	[ˈpɔkɐ]

of the tongue contours) (Albuquerque, Valente, Barros, et al., 2021; Barros, Silva, et al., 2020; Barros, Valente, et al., 2020).

Additionally, a prevention and action plan to implement during the data acquisition due to the pandemic situation was developed, in order to deal with hygiene and safety issues (see Appendix H). This contingency plan follows the Prevention and Action Plan against COVID-19 of the University of Aveiro, the guidelines of the World Health Organization and the European Centre for Disease Prevention and Control, as well as the National Contingency Plan and the Guidelines issued by the Directorate-General of Health, in the context of the prevention and control of infection by SARS-CoV-2 virus.

The participants filled in the background questionnaire and answered the screening questionnaire for COVID-19 (see Appendix H), which took approximately 15 minutes to complete. The investigator recruited and collected data for all participants, with collaboration of other researchers.

Regarding the corpus acquisition, the participants were asked to seat, facing a computer screen displaying prompts, and to wear a stabilization helmet (Articulate Instruments Ltd., 2008), in order to ensure that neither the speaker's head nor the transducer moved during the experiment. The experimental setup for data acquisition is shown in Figure 6.1.

Synchronous acquisition of US images and speech sounds was performed through Articulate Assistant Advanced software (AAA) (Articulate Assistant Ltd., 2014) and took place in a quiet room using an endocavitary probe (65EC10EA) with 90º field of view positioned under the participants' chin. The US images were collected at 60 frames per second using a Mindray DP6900 US equipment with depth set to 97 mm. The speech audio was collected with a Philips SBC ME400 microphone connected to an external sound system (UA-25 EX USB). The synchronization between the video and audio streams was performed in AAA based on the synchronization pulses introduced, during recording, by a SyncBrightUp module (Articulate Instruments Ltd., 2010; Wrench & Scobbie, 2008).

Instructions were provided prior to recording to ensure familiarity with the speech materials. Speakers were instructed to read the sentences at a comfortable pace, and during recording, stimuli that were mispronounced were repeated and the repeated recording replaced the original productions.

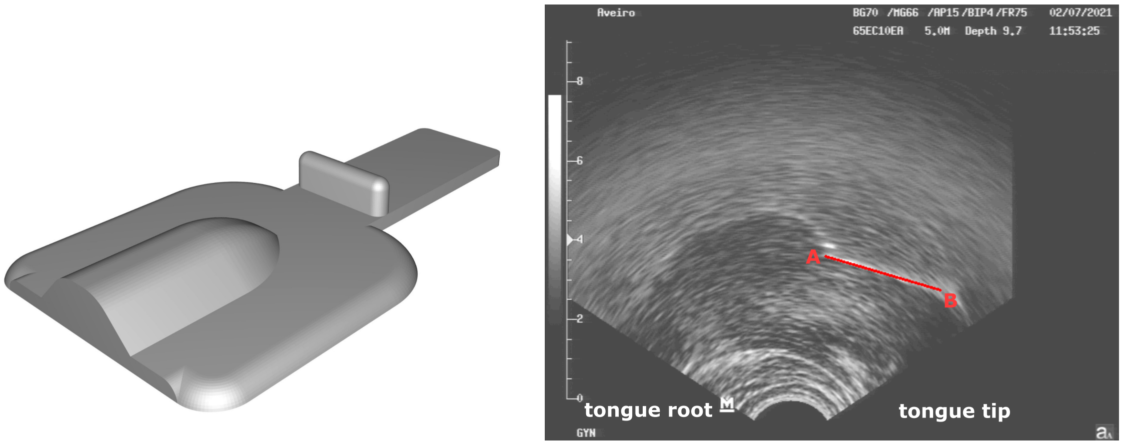
The speech material was presented in three randomized blocks (i.e., front ([i], [e], [ɛ]), central ([ɨ], [ɐ], [a]) and back ([u], [o], [ɔ]) vowels). At the start of each block a recording of the sequence [tatatata] was obtained to provide additional data for further assessing the audio-video synchronization during the data processing stage. Also, each block began and finished with the production of the sustained vowel [a] and the recording of the bite plane. The bite plane was recorded in order to image the occlusal plane, which is a reliable method for the definition of horizontal and vertical orientations in the vocal tract (Dokovova, Sabev, Scobbie, Lickley, & Cowen, 2019; Scobbie et al., 2011). That is, the speaker was asked to bite and press their tongue against a flat plastic plate, with 4 cm length, back from the upper incisor, which results in their tongue bulging upward at the back edge of the bite plate (Dokovova et al., 2019; Scobbie et al., 2011). This bite plate was designed by the authors for this experiment and is shorter than the one used in Scobbie et al. (2011), to prevent the gag reflex (see Figure 6.2**a**). The bite plate was 3D printed in polyethylene terephthalate glycol-modified (PETG) filament (natural color).

Figure 6.1: Experimental setup for recording ultrasound: the participant is seated in front of a microphone, wearing a stabilization helmet to keep the ultrasound probe positioned below the chin throughout the experiment. The prompts are presented on a computer monitor. (adapted from Albuquerque, Valente, et al. (2022))

The probe orientation was adjusted between the shadows of the mandible and hyoid bones, taking into account different anatomical characteristics of each speaker. The bite plane sequences and the contours of the sustained vowels [a] were then used to obtain the referential, for each speaker. To ensure that for each speaker there is at least one bite plane trace to obtain the referential, the probe orientation was fixed along the session.

Finally, each carrier sentence was repeated 3 times. Therefore, each speaker produced 81 individual utterances (i.e., 9 repetitions of each vowel, per context) and needed approximately 30 minutes to complete the task.

At the end of each session the ultrasound data was synchronized with the speech data by aligning the audio tone, which was triggered at the onset of each recording, and the video flash (i.e., a bright flash on the corner of the ultrasound image) (Kirkham & Nance, 2017; Wrench & Scobbie, 2008). After that, the US images and the acoustic files were exported from AAA software. This articulatory database is analyzed in Chapter 11.



(a) **(b)**

Figure 6.2: To ensure a common referential across blocks and sessions, US acquisitions were considered including a bite plate (**a**). The corresponding images were annotated (**b**) and the data informed contour adjustment (e.g., rotation) before analysis. (**a**) Bite plate; (**b**) Bite plane trace. (adapted from Albuquerque, Valente, et al. (2022))

This page intentionally contains only this sentence.

PART III

ACOUSTIC AND ARTICULATORY STUDIES

This page intentionally contains only this sentence.

AGE-RELATED ACOUSTIC CHANGES IN EUROPEAN PORTUGUESE ORAL VOWELS

The main body of this chapter is based on the study published in Journal of Voice [1].

The knowledge about the age effects in speech acoustics is still disperse and incomplete. For that, the purpose of this study is to analyze the age and gender effects on duration, F0 and formant frequencies (F1, F2 and F3) for all EP oral vowels produced by a large group of healthy speakers, which is essential to provide a more complete view of age-related changes in EP vowel acoustics.

A further original aspect of this study consists in looking at various possible dimensions of variation of a vowel system as function of age. The relationship between age, gender and VSA, which is a well established acoustic metric (Arias-Vergara, Vásquez-Correa, & Orozco-Arroyave, 2017; McCloy, Wright, & Souza, 2014; Mou et al., 2019; Neel, 2008; Roy et al., 2009), is analyzed. VSA characterize the distribution of the vowels in the acoustic space defined by F1 and F2, and it is used to model possible reduction in the articulatory capability of speakers (Arias-Vergara et al., 2017; Fougeron & Audibert, 2011). Such reduction is observed as a compression of the area of the vocal space. Additionally, to complement the results obtained with VSA, other vowel space metrics were adopted to further investigate age and gender effects. The F1RR and F2RR were selected to model possible reduction in F1 or F2, which can be interpreted as impairment in the range of high/low or front/back tongue movements (and rounding for the F2 of back vowels), respectively (Audibert, Fougeron, Gendrot, & Adda-Decker, 2015; Fougeron & Audibert, 2011; Gahl & Baayen, 2019). VAI and FCR were included to maximize sensitivity to vowel formant centralization and minimize sensitivity to interspeaker variability (Sapir, Ramig, Spielman, & Fox, 2011, 2010). The general assumption is that young speakers have a better articulation capability than older speakers, thus young adults are able to move their tongue with greater amplitudes and they are able to hold it longer in certain positions (Rastatter & Jacques, 1990; Rastatter et al., 1997).

As novelty this study considers age as a continuous variable in the analysis avoiding the effects of arbitrary age groups division. Thus, age-related changes in vowel acoustics are analyzed using multiple linear regression. Furthermore, an in-depth analysis of the age effect on each vowel was performed.

Since there is a paucity of literature on EP vowel acoustics and the available data were collected from a small number of speakers (Albuquerque et al., 2014; Escudero et al., 2009; M. R. D. Martins, 1973; Oliveira, Cunha, et al., 2012), this study also provides valuable insights to an accurate description of the behavior of each vowel.

[1] Albuquerque, L., Oliveira, C., Teixeira, A., Sa-Couto, P., & Figueiredo, D. (2020). A comprehensive analysis of age and gender effects in European Portuguese oral vowels. *Journal of Voice*, (in press).

7.1 Method

7.1.1 Speakers and speech data

The EP speech database used in this study have previously been described in Chapters 4 and 5. Briefly, vowel tokens were produced by 112 healthy native Portuguese speakers, without voice or speech problems, aged between 35 and 97 (56 men and 56 women).

The speech corpus consisted of 36 disyllabic words, with the EP oral vowels [i], [e], [ɛ], [a], [o], [ɔ], [u] in stressed position and the vowels [ɨ] and [ɐ] in unstressed position. Each participant produced 12 repetitions of each vowel, in a total of 108 productions by speaker (112 participants x 36 words x 3 repetitions = 12 096 recordings).

For more details about the study design, recording protocol, segmentation and speakers characterization see Chapters 4 and 5.

7.1.2 Acoustic measurements

The acoustic parameters F0, F1, F2, F3 and vowel duration were extracted as described in Section 5.3.1 (Segmental Parameters). In order to characterize the vowel space, the acoustic metrics VSA, F1RR, F2RR, VAI and FCR were obtained.

The VSA is defined by the polygon area based on the mean value for each stressed oral vowel, adapted from Arias-Vergara et al. (2017), McCloy et al. (2014) and Neel (2008).

The F1RR is defined as the ratio of the F1 of the low vowel [a] and the (geometric) average F1 of the high vowels [i] and [u] by speaker (Audibert et al., 2015; Escudero et al., 2009; Fougeron & Audibert, 2011). The F2RR is computed as the ratio of the F2 of vowel [i] and the F2 of the vowel [u] for each speaker (Audibert et al., 2015; Escudero et al., 2009; Fougeron & Audibert, 2011; Sapir et al., 2010).

The VAI is calculated using the formula:

$$\text{VAI} = (\text{F2[i]} + \text{F1[a]})/(\text{F2[u]} + \text{F2[a]} + \text{F1[u]} + \text{F1[i]}), \tag{7.1}$$

and its inverse, the FCR, is calculated as:

$$\text{FCR} = (\text{F2[u]} + \text{F2[a]} + \text{F1[u]} + \text{F1[i]})/(\text{F2[i]} + \text{F1[a]}) \tag{7.2}$$

(Roy et al., 2009; Sapir et al., 2011, 2010). Note that the F1 and F2 coordinates of the EP corner vowels [a], [i] and [u] were used to calculate the VAI and FCR metrics, so the FCR should increase with centralization and decrease with vowel expansion, and the opposite for VAI (Sapir et al., 2011, 2010).

7.1.3 Statistical analysis

The statistical analysis was conducted with the SPSS software package (SPSS 25.0 - SPSS Inc., Chicago, IL, USA). The values of F0, F1, F2, F3 and duration were computed for all productions, and subsequently, the median of repetitions was performed for each vowel and speaker.

For each vowel and acoustic parameter (duration, F0, F1, F2 and F3), a multiple linear regression was conducted with the following explanatory variables: age (continuous), gender (male: reference group, female), and the interaction between age and gender. The model presented by the software considered age and gender (female) redundant (presenting instead the interactions "male*age" and "female*age") and no values were presented for those (independent) variables. Also, a multiple linear regression was conducted with the same explanatory variables for VSA, F1RR, F2RR, VAI, FCR and for mean values of all vowels by acoustic parameter. The regression coefficients and the correspondent 95% confidence interval (95%CI) were calculated. The residuals Normality was tested (Kolmogorov-Smirnov Test) and verified with the visual inspection of the Q-Q plot.

7.2 Results

This section presents the detailed results of the acoustic measurements and statistical analysis aimed at investigating differences by age and gender for duration, F0 and formant frequencies of all vowels. To avoid effects of arbitrary age groups division, correlation and regresssion analyses for all acoustic parameters were performed.

7.2.1 Vowel duration increased with age

Scatterplot and regression results for the duration are presented in Figure 7.1, and show an increase of mean duration for all vowels with age, for both genders. For females, duration increased from approximately 100 ms to more than 140 ms between the ages 35 and 100; the increase was lower for males, only reaching 130 ms at the age of 100.

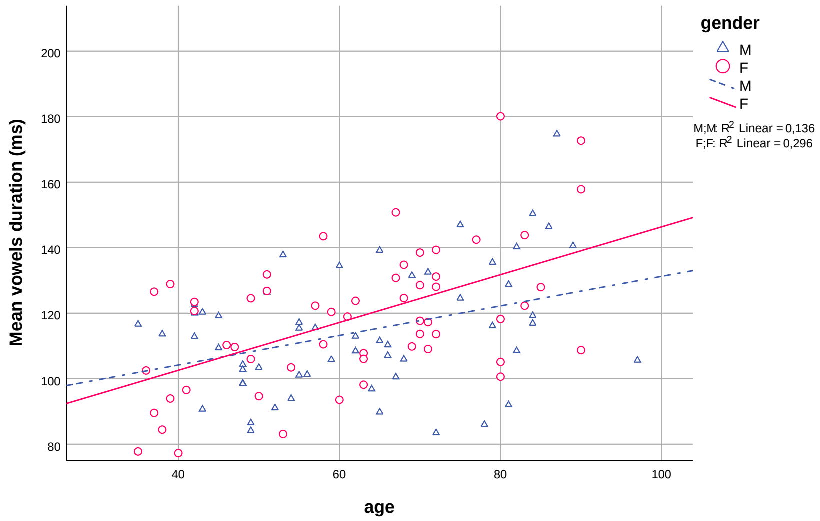


Figure 7.1: Scatterplot and regression lines for vowels duration by age and gender. Each symbol corresponds to one speaker. Solid line and circles: females; dashed line and triangles: males. (adapted from Albuquerque et al. (2020))

The multiple linear regression revealed a significant effect of age in both genders, for most vowels and for the mean of all vowels (males: $B = 0.451$; $p = 0.004$; 95%CI= [0.144; 0.759]; females: $B = 0.730$; $p < 0.001$; 95%CI= [0.427; 1.033]). Only vowel [ɨ] in both men ($B = 0.138$; $p = 0.258$; 95%CI=[−0.103; 0.379]) and women ($B = 0.183$; $p = 0.129$; 95%CI=[−0.054; 0.421]) did not seem to be significantly affected by age. There was not a significant effect of gender, with men (114.4 ms ± 19.0) and women (118.3 ms ± 21.2) producing vowels with similar mean duration ($B = 12.754$; $p = 0.362$; 95%CI=[−14.854; 40.362]).

The pattern of vowel duration was as follows: [ɐ] (60.4 ms ± 9.8) < [ɨ] (75.0 ms ± 14.1) < [i] (119.4 ms ± 26.5) < [u] (122.4 ms ± 29.3) < [ɛ] (126.0 ms ± 24.7) < [o] (126.7 ms ± 25.6) < [ɔ] (130.4 ms ± 27.1) < [a] (143.2 ms ± 25.3) < [e] (143.5 ms ± 24.6) .

7.2.2 Age effects in F0 were gender dependent

The scatterplot of F0 mean vowels is presented in Figure 7.2 and shows that mean F0 tended to decrease with age in women and slightly increase in men. Regression lines indicated a decrease for females of about 25 Hz between the ages 35 and 100, and an increase around 10 Hz for males between the same ages.

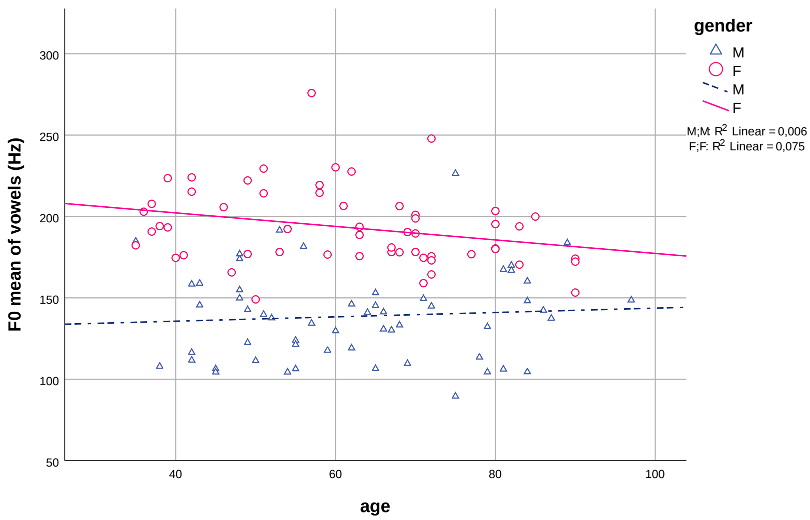


Figure 7.2: Scatterplot and regression lines for mean F0 by age and gender. Each symbol corresponds to one speaker. Solid line and circles: females; dashed line and triangles: males. (adapted from Albuquerque et al. (2020))

Regression model revealed a main effect of gender ($B = -88.482$; $p < 0.001$; 95%CI=[−128.000; −48.964]), since male speakers had significantly lower F0 (138.7 Hz ± 27.6) compared to female speakers (193.3 Hz ± 23.9), as expected. The effects of age in both genders were not significant for the majority of the vowels, except for the unstressed vowels in females ([ɨ]: $B = -0.766$; $p = 0.003$; 95%CI=[−1.271; −0.262]; [ɐ]: $B = -0.954$; $p < 0.001$; 95%CI=[−1.434; −0.475]). In these vowels F0 decreased very sharply with age.

As illustrated in Figure 7.3, which was drawn using equations of linear regression (of all vowels by gender and F0) replacing the variable age for 35 and 100, F0 frequencies of all vowels tended to decrease in women (mainly in unstressed vowels) and to slightly increase in men (except the unstressed vowels), with aging. Therefore, F0 tended to approach between genders as age increased.

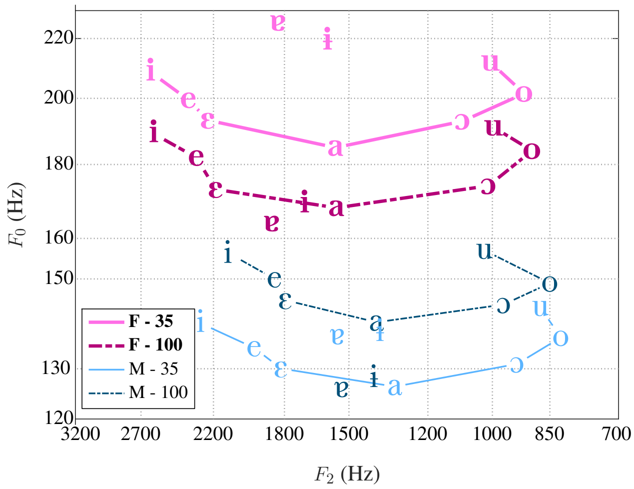


Figure 7.3: Mean value of F0 as a function of vowel and age. Top: women; bottom: men. Solid lines: 35 years; dashed lines: 100 years. This figure was drawn using equations of linear regression (of each vowel by gender) replacing the variable age by 35 and 100 (as an approximation to the age of the oldest speaker of the sample). (adapted from Albuquerque et al. (2020))

Additionally, as it can be seen in Figures 7.3 and 7.4, there were F0 differences between vowels of different phonological heights. Central low vowel [a] displays the lowest mean value of F0 (155.2 Hz ± 33.8), followed by the vowels ([ɛ] (160.7 Hz ± 35.8) < [ɔ] (160.8 Hz ± 35.9) < [ɨ] (166.1 Hz ± 44.7) < [e] (166.5 Hz ± 38.0) < [ɐ] (166.7 Hz ± 45.3) < [o]) (168.3 Hz ± 38.5)) and finally, the vowels ([i] (173.5 Hz ± 39.7) < [u] (176.0 Hz ± 40.3)).

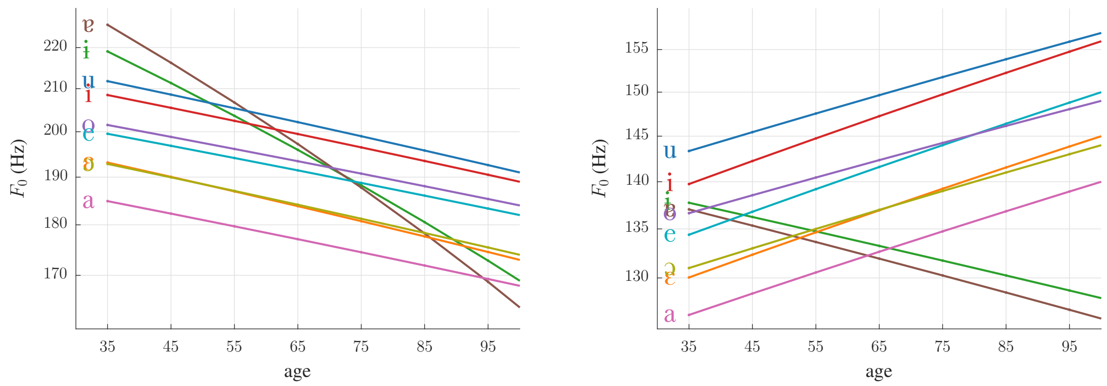


Figure 7.4: Linear regression of F0 of all vowels as function of age by gender. Left graph: females; right graph: males.

In general, back vowels have a higher F0 than front vowels, but the data suggests that this effect tends to disappear with age, mainly in males (see Figure 7.4).

7.2.3 Age effects in formant frequencies were vowel and gender dependent

As in previous sections, analysis of vowel formants start by showing scatterplots and regression of mean frequencies (Figures 7.5, 7.6 and 7.7). Vowel space based on regression results by gender and age is presented in Figure 7.8. Due to the complexity of the results, the multiple linear regression coefficients are displayed in Table 7.1.

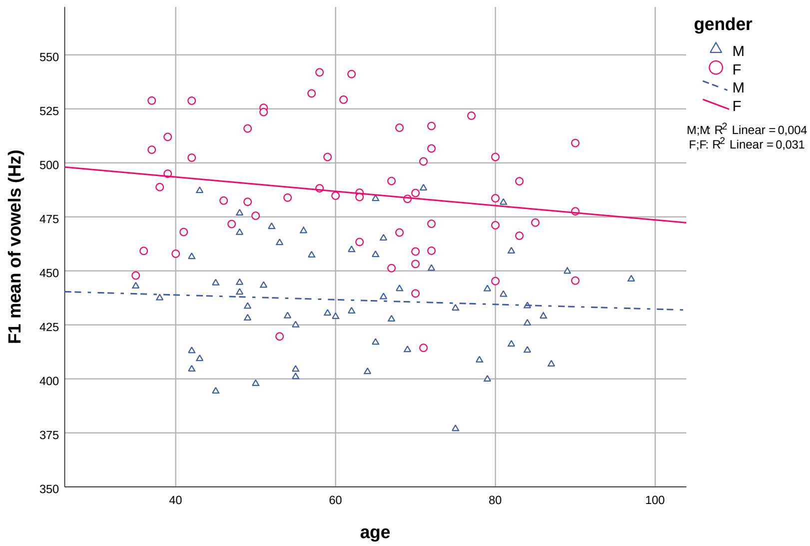


Figure 7.5: Scatterplot and regression lines for mean F1 by age and gender. Each symbol corresponds to one speaker. Solid line and circles: females; dashed line and triangles: males. (adapted from Albuquerque et al. (2020))

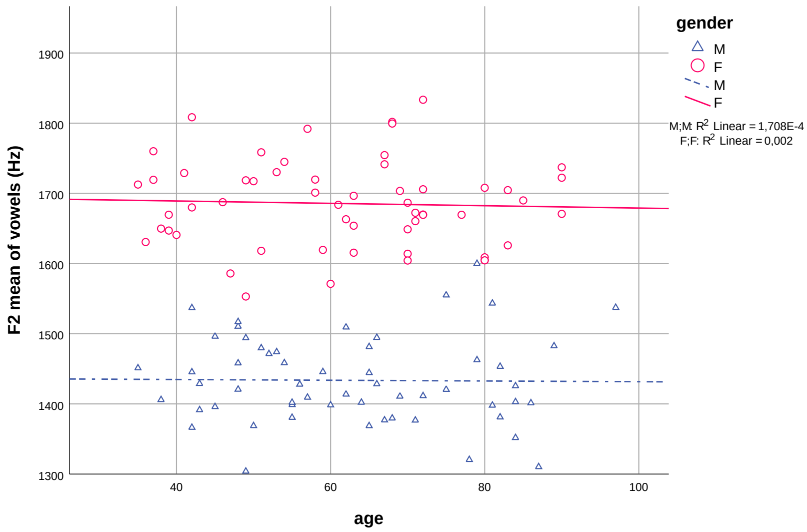


Figure 7.6: Scatterplot and regression lines for mean F2 by age and gender. Each symbol corresponds to one speaker. Solid line and circles: females; dashed line and triangles: males. (adapted from Albuquerque et al. (2020))

Figure 7.5 shows that mean F1 tended to decrease with age in both genders, mainly in women. The results of multiple linear regression coefficients (Table 7.1) revealed significant differences between genders for all vowels (except for central vowels ([ɨ], [ɐ] and [a]) and [ɛ]), with women presenting

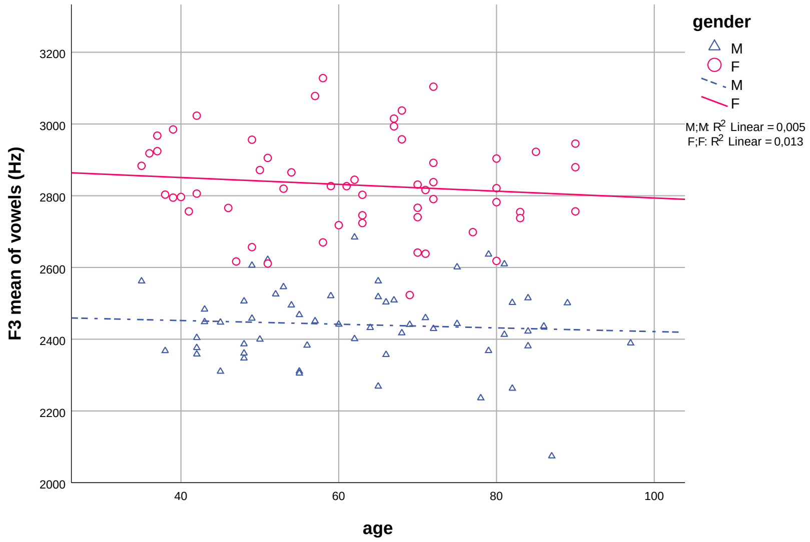


Figure 7.7: Scatterplot and regression lines for mean F3 by age and gender. Each symbol corresponds to one speaker. Solid line and circles: females; dashed line and triangles: males. (adapted from Albuquerque et al. (2020))

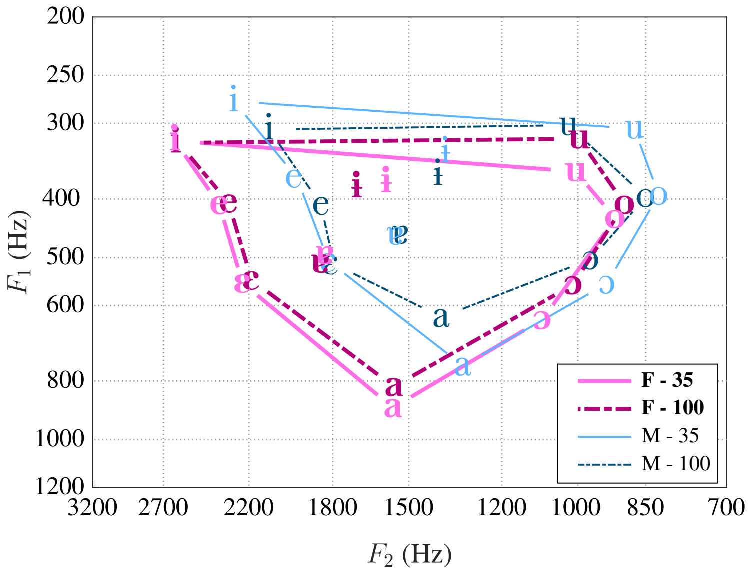


Figure 7.8: Vowel Space for men and women as a function of age. Bold lines and symbols: women; non bold lines and symbols: men. Solid lines: 35 years; dashed lines: 100 years. This figure was drawn using equations of linear regression (of each vowel by gender for F1 and F2) replacing the variable age for 35 and 100 (as an approximation to the age of the oldest speaker of the sample). (adapted from Albuquerque et al. (2020))

significantly higher mean F1 (486.3 ms ± 29.7) than men (436.4 ms ± 25.9). There was a significant age effect: in males for vowels [a] (Figure 7.9) and [e]; and in females for vowels [ɔ] and [u].

Summing up, as illustrated in Figure 7.8, F1 decreased with age, especially for vowels [a] and [ɔ], but increased for vowels [e], [i] and [ɨ] in males. In females, F1 decreased with age, especially for vowels [u], [ɔ] and [a].

Table 7.1: Results of multiple linear regression: the effect of gender and gender*age interaction on F1, F2 and F3 values by vowel. Grey cells represent significant results (p<0.05). (adapted from Albuquerque et al. (2020))

	Vowel	Intercept			Gender (Male)			Male * Age			Female * Age		
		B	95% CI		B	95% CI		B	95% CI		B	95% CI	
F1	ɨ	371.54	329.35	413.73	-52.17	-112.77	8.42	0.46	-0.21	1.14	0.11	-0.55	0.78
	ɐ	484.63	433.43	535.82	-25.67	-99.20	47.87	-0.03	-0.85	0.79	0.25	-0.55	1.06
	a	921.76	832.36	1011.17	-99.06	-227.47	29.36	-1.96	-3.39	-0.53	-1.05	-2.45	0.36
	e	409.51	375.87	443.14	-62.75	-111.06	-14.45	0.62	0.08	1.16	-0.04	-0.57	0.49
	ɛ	554.03	509.26	598.81	-49.76	-114.07	14.56	0.08	-0.63	0.80	-0.10	-0.80	0.61
	i	318.92	286.36	351.48	-59.60	-106.37	-12.84	0.48	-0.04	1.00	0.04	-0.47	0.55
	o	442.72	408.10	477.35	-50.73	-100.46	-1.00	0.05	-0.51	0.60	-0.35	-0.90	0.20
	ɔ	674.54	618.92	730.16	-96.78	-176.66	-16.89	-0.75	-1.64	0.14	-1.21	-2.08	-0.33
	u	382.67	349.98	415.36	-74.90	-121.85	-27.95	0.06	-0.46	0.58	-0.65	-1.16	-0.14
	Mean	506.70	476.77	536.64	-63.49	-106.49	-20.49	-0.11	-0.59	0.37	-0.33	-0.80	0.14
F2	ɨ	1523.01	1390.08	1655.93	-165.49	-356.40	25.42	0.39	-1.74	2.51	1.79	-0.30	3.88
	ɐ	1815.24	1710.26	1920.22	-254.28	-405.06	-103.50	-0.30	-1.97	1.38	0.31	-1.35	1.96
	a	1563.46	1478.31	1648.62	-287.37	-409.67	-165.06	1.07	-0.29	2.43	-0.06	-1.41	1.28
	e	2392.12	2239.69	2544.55	-373.59	-592.51	-154.66	-1.69	-4.12	0.75	-0.81	-3.21	1.59
	ɛ	2255.71	2121.54	2389.87	-430.60	-623.30	-237.90	-0.35	-2.50	1.79	-0.74	-2.85	1.38
	i	2642.45	2487.84	2797.06	-264.11	-486.18	-42.05	-2.85	-5.32	-0.38	-0.37	-2.80	2.07
	o	924.77	859.08	990.46	-114.96	-209.31	-20.61	0.40	-0.65	1.45	-0.31	-1.34	0.73
	ɔ	1133.93	1071.55	1196.31	-220.95	-310.55	-131.36	0.56	-0.44	1.56	-1.19	-2.18	-0.21
	u	1012.17	882.36	1141.98	-219.64	-406.08	-33.20	2.30	0.23	4.38	-0.14	-2.18	1.91
	Mean	1695.87	1628.39	1763.36	-259.00	-355.92	-162.07	-0.05	-1.13	1.03	-0.17	-1.23	0.89
F3	ɨ	2875.24	2713.98	3036.49	-271.48	-503.08	-39.88	-1.89	-4.47	0.69	0.47	-2.07	3.01
	ɐ	2805.99	2644.83	2967.15	-366.92	-598.39	-135.46	-1.40	-3.98	1.17	-0.74	-3.28	1.80
	a	2561.37	2354.48	2768.25	-315.25	-612.39	-18.11	1.74	-1.56	5.05	1.27	-1.99	4.53
	e	2912.49	2760.88	3064.10	-424.42	-642.17	-206.67	-0.11	-2.53	2.31	-0.02	-2.41	2.36
	ɛ	2903.41	2736.52	3070.30	-502.15	-741.85	-262.44	0.82	-1.85	3.49	-1.15	-3.77	1.48
	i	3257.24	3024.54	3489.94	-313.60	-647.82	20.62	-3.05	-6.77	0.67	-2.32	-5.98	1.35
	o	2913.92	2724.79	3103.05	-484.69	-756.33	-213.05	-0.37	-3.39	2.66	-0.98	-3.96	2.00
	ɔ	2766.86	2554.54	2979.18	-447.83	-752.78	-142.88	0.62	-2.78	4.01	-0.69	-4.03	2.65
	u	3002.76	2819.95	3185.58	-619.73	-882.30	-357.16	-0.97	-3.90	1.95	-4.40	-7.28	-1.52
	Mean	2888.81	2757.77	3019.84	-416.23	-604.43	-228.03	-0.51	-2.61	1.58	-0.95	-3.01	1.11

B = Linear Coefficient

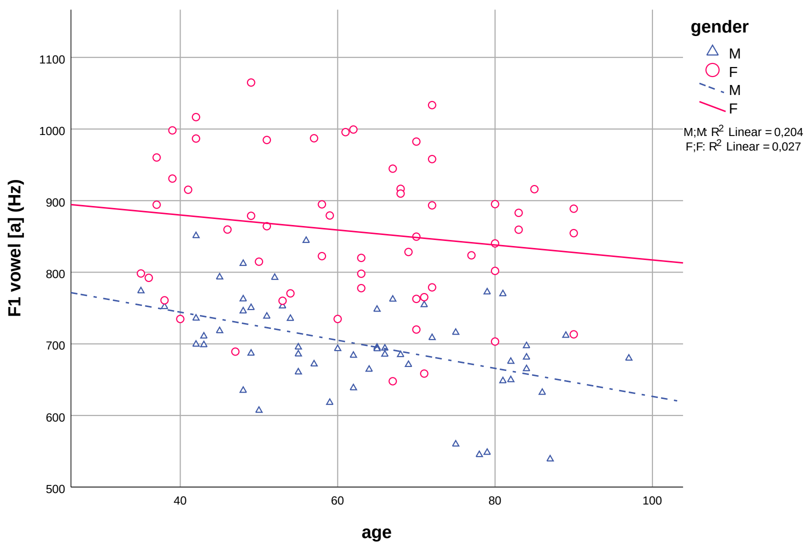


Figure 7.9: Scatterplot of F1 as function of age and gender for vowel [a] with superimposed linear regression results. Each symbol corresponds to one speaker. Solid line and circles: females; dashed line and triangles: males. (adapted from Albuquerque et al. (2020))

Concerning the vowel effect, the main determiner of F1 is the phonological vowel height: the low vowel [a] has the highest F1 (778.6 Hz ± 116.5), followed by the lower mid vowels ([ɔ] (565.4 Hz ± 63.8) > [ɛ] (528.7 Hz ± 45.5) > [ɐ] (478.7 Hz ± 51.7)), then the higher mid vowels ([o] (408.1 Hz ± 34.6) > [e] (396.5 Hz ± 33.4)), and finally the high vowels ([ɨ] (363.4 Hz ± 41.9) > [u] (327.2 Hz ± 34.6) > [i] (305.5 Hz ± 34.3)) which have the lowest F1.

The mean values of F1 and the regression lines by vowel (Figure 7.10) show that each back vowel has a higher F1 than its front counterpart in younger ages, but these differences tended to disappear with age in both genders. The approximation of F1 values of the back-front vowel pairs occurred due to a tendency of F1 decrease on back vowels in females, while in males a tendency of an increase of the F1 of front vowels was observed.

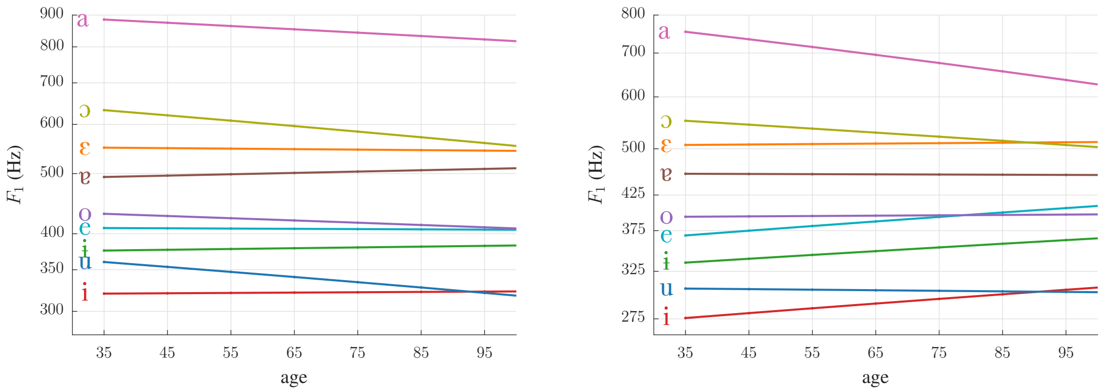


Figure 7.10: Linear regression of F1 of all vowels as function of age by gender. Left graph: females; right graph: males.

As shown in Figure 7.6, mean F2 did not reveal remarkable changes with age. The statistical analysis revealed a main effect of gender on F2 (see Table 7.1): women's mean F2 frequencies (1685.5 Hz ± 62.3) were significantly higher than those of men (1433.6 Hz ± 62.1). Only vowel [ɨ] did not present significant differences between genders.

A reliable effect of age was found for some vowels depending on gender. Considerable decrease in F2 was found for males in vowel [i] and an opposite tendency was observed in [u] (see Figure 7.11). Female [ɔ] displayed an F2 decrease with aging.

The mean value of F2 was higher for vowel [i] (2409.9 Hz ± 256.2) followed by [e] (2127.6 Hz ± 257.9) , [ɛ] (2006.7 Hz ± 238.9) , [ɐ] (1688.3 Hz ± 175.4) , [ɨ] (1507.4 Hz ± 176.8) , [a] (1451.3 Hz ± 134.4) , [ɔ] (1004.2 Hz ± 81.7) , [u] (970.3 Hz ± 126.3) and [o] (870.3 Hz ± 70.2).

Mean F3 of the vowels (Figure 7.7) tended to slightly decrease with age in both genders, mainly in women.

As seen in Table 7.1, there was a significant effect of gender for F3, with males (2440.5 ms ± 109.3) to have lower F3 mean values than females (2830.3 ms ± 132.3). In addition, there was no significant age effect in men and women, except for female [u], that decreased sharply with age.

The mean value of F3 was higher for vowel [i] (2933.5 Hz ± 283.6) followed by [e] (2696.1 Hz ± 256.8) , [ɨ] (2694.8 Hz ± 257.9) , [ɛ] (2642.8 Hz ± 245.3) , [o] (2629.9 Hz ± 284.0) , [ɐ] (2555.9 Hz ± 253.8) , [ɔ] (2541.0 Hz ± 268.2) , [u] (2527.1 Hz ± 270.2) and [a] (2497.4 Hz ± 238.6).

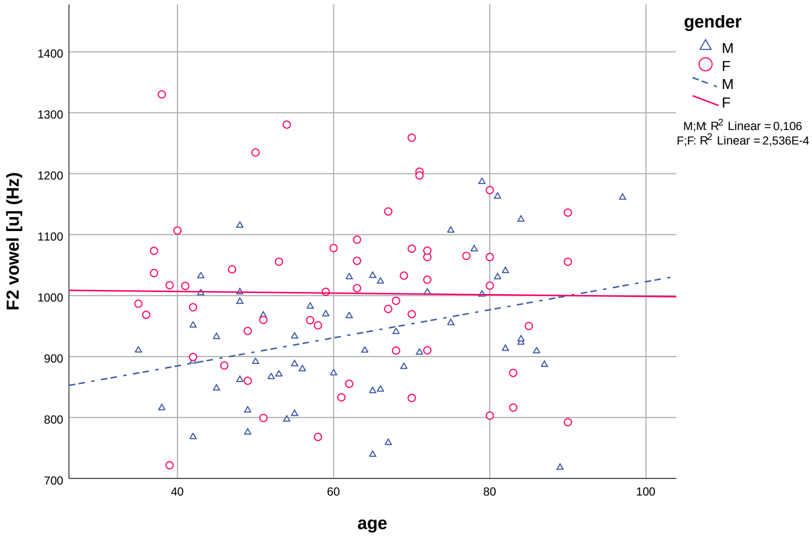


Figure 7.11: Scatterplot of F2 as function of age and gender for vowel [u] with superimposed linear regression results. Each symbol corresponds to one speaker. Solid line and circles: females; dashed line and triangles: males. (adapted from Albuquerque et al. (2020))

7.2.4 Age and gender effects on acoustic vowel space

Changes in vowel space size were computed in order to track the relationship between speaker age and vowel centralization or expansion.

For VSA (see Figure 7.12), the regression lines show a slight decrease with age in both genders, especially in men. However, the statistical model did not confirm the age tendencies reported (males: $B = -1676.398$; $p = 0.083$; 95%CI=[$-3573.152; 220.355$]; females: $B = -627.521$; $p = 0.507$; 95%CI=[$-2495.767; 1240.725$]), and also did not reveal a main effect of gender ($B = -122174.600$; $p = 0.158$; 95%CI=[$-292591.159; 48241.959$]).

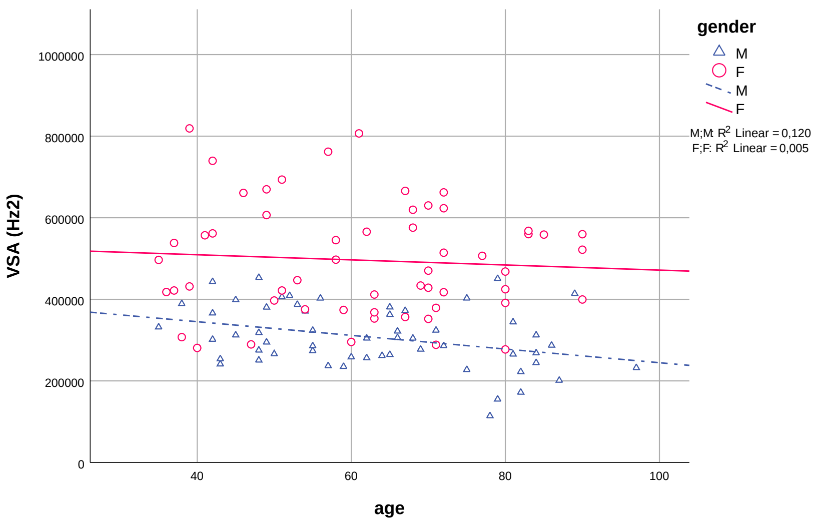


Figure 7.12: Scatterplot and regression lines for VSA by age and gender. Each symbol corresponds to one speaker. Solid line and circles: females; dashed line and triangles: males.

The scatterplot of the F1RR is presented in Figure 7.13. The regression lines show a decrease with age in both genders, mainly in males, whose F1RR decreased from 2.6 to 2.0 as age increased. The multiple linear regression results revealed that for F1RR only the age effect on males was significant ($B = -0.009$; $p = 0.006$; 95%CI=$[-0.015; -0.003]$). Moreover, the average F1RR of women was 2.609 (SD=0.422) and 2.345 (SD=0.307) for men. The female F1 space was therefore 2.609/2.345=1.135 times bigger than the male F1 space, although statistical model did not reveal a main effect of gender ($B = 0.220$; $p = 0.433$; 95%CI=$[-0.335; 0.775]$).

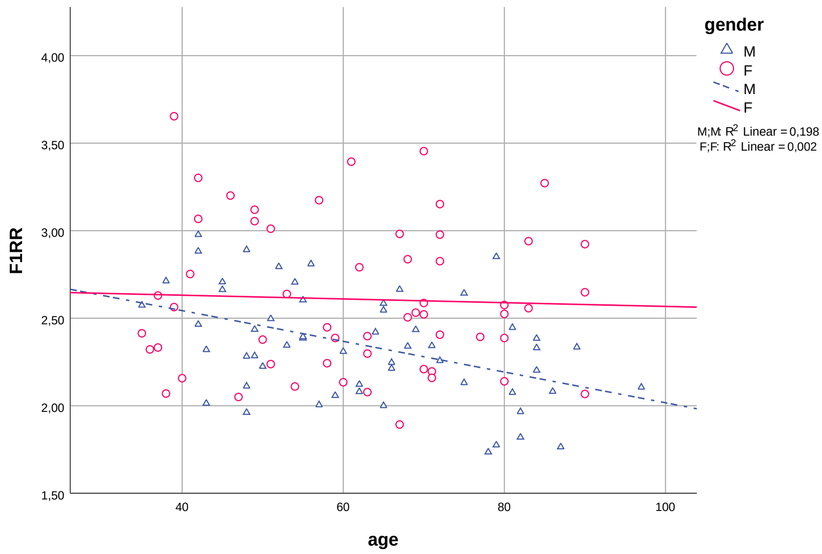


Figure 7.13: Scatterplot and regression lines for F1RR by age and gender. Each symbol corresponds to one speaker. Solid line and circles: females; dashed line and triangles: males. (adapted from Albuquerque et al. (2020))

Figure 7.14 presents the mean F2RR and indicates a decrease with age only in males (F2RR decreased around 0.5 points between the ages 35 and 100). The effect of age and gender on F2RR was also analyzed, and as with F1RR, the statistical analysis only revealed a main effect of age in males ($B = -0.008$; $p = 0.016$; 95%CI=$[-0.015; -0.002]$). Similar to F1, the size of the F2 space was higher for women (2.662) than for men (2.382), i.e. the female F2 space was therefore 1.118 times bigger than the male F2 space. Nonetheless, the model did not include a main effect of gender ($B = 0.204$; $p = 0.490$; 95%CI=$[-0.380; 0.787]$).

VAI (see Figure 7.15) and FCR were also analyzed. The regression lines of VAI show a decrease of approximately 18% between the ages of 35 and 100 for males. As expected, for males, the opposite trend was observed for FCR. For females, both parameters remained stable with age. The multiple linear regression results revealed that, for both parameters, the age effect is only significant for men (FCR: $B = 0.003$; $p < 0.001$; 95%CI=$[0.002; 0.005]$; VAI: $B = -0.003$; $p = 0.001$; 95%CI=$[-0.005; -0.001]$). As expected (they act as interspeaker normalization), the statistical model did not reveal a main effect of gender (FCR: $B = -0.127$; $p = 0.067$; 95%CI=$[-0.263; 0.009]$; VAI: $B = -0.099$; $p = 0.185$; 95%CI= $[-0.048; 0.247]$).

Additionally, Figure 7.8 allows to verify that males and females showed different tendencies with age, which is reflected in differences in vowel space sizes.

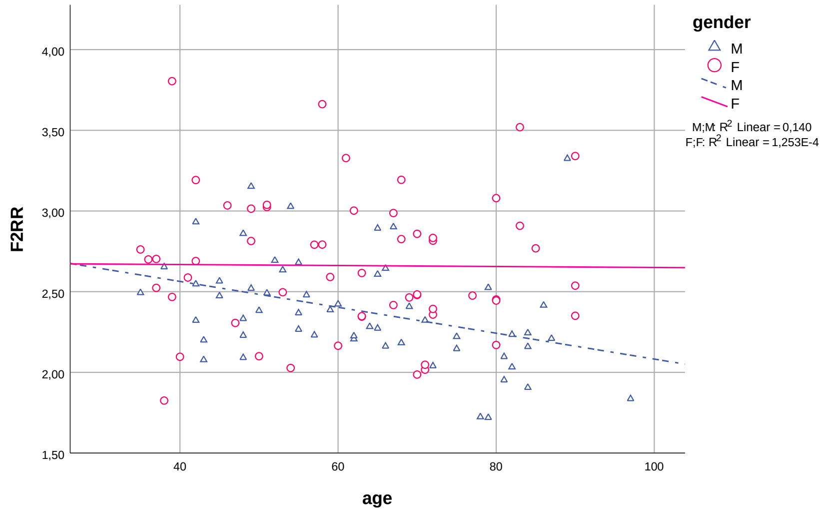


Figure 7.14: Scatterplot and regression lines for F2RR by age and gender. Each symbol corresponds to one speaker. Solid line and circles: females; dashed line and triangles: males. (adapted from Albuquerque et al. (2020))

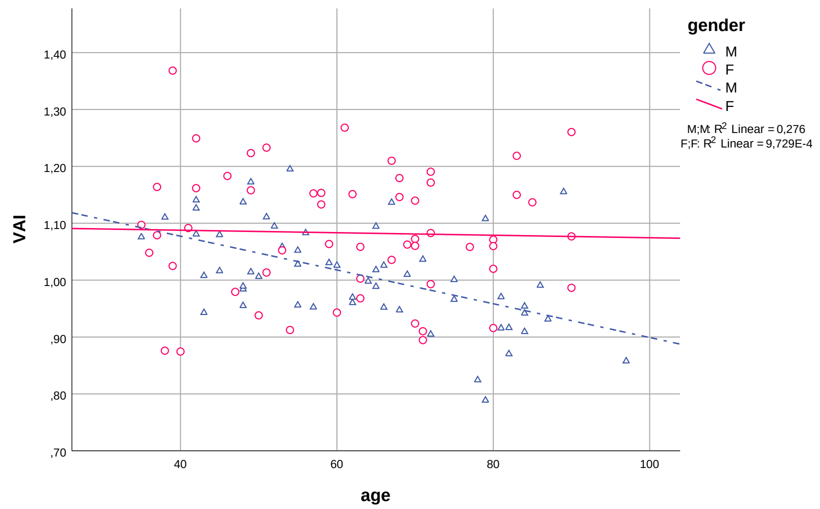


Figure 7.15: Scatterplot and regression lines for VAI by age and gender. Each symbol corresponds to one speaker. Solid line and circles: females; dashed line and triangles: males. (adapted from Albuquerque et al. (2020))

7.3 Discussion

This study contributes to increase knowledge on EP aging speech, providing an acoustical perspective of the effects of age (as a continuous variable) in all oral vowels of the EP for several acoustic parameters (duration, F0, F1, F2, F3, VSA, F1RR, F2RR, VAI and FCR). Firstly, vowels' duration increased with age. Secondly, a tendency for f0 to decrease in women and to slightly increase in men was observed. Thirdly, F1 and F2 space underwent a reduction in males with aging. Finally, the frequencies of F3 were essentially unchanged with age.

The results obtained are in general in line with previous research. However, some features, specially

F0, did not yield as much age-related variation as reported in previous studies (Eichhorn et al., 2018; Pellegrini et al., 2013; Torre III & Barlow, 2009). It should be noted that these studies used different methodologies and different criteria for participant selection (Eichhorn et al., 2018; Kent & Vorperian, 2018; Lutzross, Schuerman, Sprouse, & Gahl, 2013), and in this sense the differences in results are not surprising.

The age-related changes on acoustical parameters are discussed in detail further below.

7.3.1 Vowels duration

As in most studies (Albuquerque et al., 2014; Benjamin, 1982; Fletcher et al., 2015; Jacewicz et al., 2009; Pellegrini et al., 2013; Schötz, 2006; Smith et al., 1987), vowels' duration was the acoustic parameter mostly affected by aging. This probably occurs as a consequence of the decrease in the speech rate (Linville, 2001; Schötz, 2006; Steffens, 2011) and seems to be related to the neuromuscular slowing, altered nerve supply, respiratory changes, increased cautiousness and to the adjustment by older speakers of their tempo to maintain speech fluency (Linville, 2001; Schötz, 2006). Additionally, Fletcher et al. (2015) suggested that the lengthening of vowels with age may assist older speakers in maintaining the articulatory precision (i.e., maintain acoustically distinct vowels).

As expected, the unstressed vowels presented the shortest vowel duration. After the vowel [ɐ], the unstressed vowel [ɨ] had the lowest duration and tended to be deleted (26,7%) (Fikkert, 2005; Pellegrini et al., 2013). The deletion of unstressed vowels, especially of [ɨ], has been reported for many languages and also for EP (Fikkert, 2005; Mateus & D'Andrade, 2000; D. J. Silva, 1994, 1998; Veloso, 2007, 2010). At the same time, [ɨ] duration remains almost unchanged with age.

7.3.2 Fundamental frequency

The results of this study give additional support that age related changes in F0 are gender dependent, which leads to an approximation of F0 values between genders as age increases. As in the current study, most of the literature for other languages reported a lowering of F0 with age for women, and a raising of F0 for men (not always significant) (Eichhorn et al., 2018; Nishio & Niimi, 2008; Reubold et al., 2010; Schötz, 2006; Sebastian et al., 2012; Torre III & Barlow, 2009; Tykalova et al., 2020). For the EP, previous studies were consistent with the decrease of F0 in females (Guimarães & Abberton, 2005; Pellegrini et al., 2013), whereas in males no significant age-related changes in F0 were observed (Albuquerque et al., 2014; Guimarães & Abberton, 2005; Pellegrini et al., 2013).

It has been suggested that the F0 drop in women with age results from the increase in vocal fold mass due to hormonal changes that occur during menopause (Ferrand, 2002; Higgins & Saxman, 1991; Linville, 2001; Ma & Love, 2010; Mautner, 2011; Pontes et al., 2005; Sataloff et al., 1997). The raise in F0 in men after middle age has been attributed to reduced vocal fold mass and stiffness of vocal fold tissues due to aging-induced atrophy of the internal thyroarytenoid (Higgins & Saxman, 1991; Hollien & Shipp, 1972; Linville, 2001; Nishio & Niimi, 2008; Sebastian et al., 2012).

Additionally, it is important to mention that unstressed vowels behaved differently from stressed vowels with age. So, in unstressed vowels, F0 tended to decrease in both genders (although only statistically significant for females) and presented different values than expected. In other words, [ɨ] and [ɐ] F0 tended to be lower than the F0 of vowel [a] with age in both genders. This finding raises questions about the usual physiological explanation for a rise of F0 in older men (Higgins & Saxman, 1991; Hollien & Shipp, 1972; Linville, 2001; Nishio & Niimi, 2008; Sebastian et al., 2012), that is, it remains unclear why a reduced mass and/or stiffness of vocal folds should affect only stressed vowels, whereas unstressed vowels show quite an opposite pattern.

Furthermore, concerning the universal and Portuguese specific principles of F0, it was observed that: women have a higher average F0 than men (due to the well-known differences in vocal fold length and size between genders (Adank et al., 2004; Escudero et al., 2009; Peterson & Barney, 1952)); high vowels have a higher F0 than low vowels (which is related to the intrinsic F0 effect of vowels (Kent & Read, 2002; Whalen & Levitt, 1995)); and back vowels have a higher F0 than front vowels (previously reported for Portuguese (Escudero et al., 2009; M. R. D. Martins, 1973) and also for other languages (Whalen & Levitt, 1995)). The cause of the F0 differences between front and back vowels remains unknown. However, in this study this distinction tends to disappear with age among men, which may be related to a lower articulatory precision.

7.3.3 Formant frequencies

A general lowering of F1 and F2 frequencies for women in all stressed vowels (although significant differences occurred only for F1[u], F1[ɔ] and F2[ɔ]) was observed; as men showed: 1) a decrease in F1 for low vowels (especially in [a]) and an increase in high vowels (mainly in [i]); 2) an F2 decrease with age for [i], and a raising of F2 for [u], which suggests that only older men showed formant frequency evidence of vowel centralization, reported in previous researches (Mertens et al., 2020; Rastatter & Jacques, 1990; Rastatter et al., 1997; Schötz, 2006; Torre III & Barlow, 2009).

For EP, an opposite tendency was observed in Pellegrini et al. (2013) and Rato et al. (2017). Pellegrini et al. (2013) have shown a greater centralization of the vowel space in younger speakers (males and females aged 19-28). Rato et al. (2017) suggested that the wider vowel space in older speakers may be due to gradual changes in the contemporary speech of Braga. However, the present study does not cover the same age ranges, for that it is difficult to make a comparison between the studies.

Although the VSA did not show centralization for both genders, all the other vowel space ratios (F1RR, F2RR, VAI and FCR) indicated significant changes consistent with the centralization of the vowel space for male speakers. These results corroborate the main hypothesis that young males have a better articulation precision than older males (Arias-Vergara et al., 2017; Fougeron & Audibert, 2011; Linville, 2001). Also, F2RR reflect restricted movements of the tongue in the anterior–posterior direction and restricted movements of the lips (rounding for [u] and retraction for [i]) (Sapir et al., 2010). For example, an increase in F2 can be caused by a more anterior tongue position, but also by a decrease in lip rounding or tongue body shape (Wieling et al., 2016). The present results tend to confirm that the F1RR, F2RR, VAI and FCR metrics are more sensitive to mild vowel articulation changes with age than VSA (Caverlé & Vogel, 2020).

Several explanations have been advanced to account for age-related changes in vowel formant frequencies (Eichhorn et al., 2018; Kent & Vorperian, 2018; Linville & Rens, 2001; Mautner, 2011;

Xue & Hao, 2003), like altered dimensions of the back cavity (Scukanec et al., 1991; Xue & Hao, 2003), changes of the shape of the oral cavity (loss of teeth and the introduction of dentures) (Mautner, 2011; Xue & Hao, 2003), diachronic or intergenerational phonetic change (Fox & Jacewicz, 2010) or slower tongue movements and loss of tongue strength (Linville, 2001; P. J. Watson & Munson, 2007).

Additionally, as in Schötz (2006) and Eichhorn et al. (2018), the frequencies of F3 were essentially unchanged with age. This result does not support the idea of vocal tract lengthening in older age reported in previous studies (Decoster & Debruyne, 1999; Endres, Bambach, & Flösser, 1971; Harrington et al., 2007; Linville & Rens, 2001; P. J. Watson & Munson, 2007; Xue & Hao, 2003). So, the lack of an aging effect on the F3 indicates that any changes found for F1 and F2 are related to specific articulatory effects (Eichhorn et al., 2018). And also, this claim could be corroborated by the findings about males' patterns of F0 change in stressed vs. unstressed vowels.

Finally, the present study supports several general properties of Portuguese vowels (Escudero et al., 2009), which are in common with other languages: women tend to have higher formant values than men (which is well understood in terms of the differences in vocal tract sizes between genders (Escudero et al., 2009; Peterson & Barney, 1952)); high vowels have a higher F1 than low vowels; back vowels present higher F1 than their front counterparts; EP contrasts front and back vowels, as back vowels present lower F2 than their front counterparts. It is interesting to note that mostly the differences of F1 between front and back vowels decrease with aging (as seen in Figure 7.10)). Thus, the older adults present a lower front-back distinction compared to younger ones.

Furthermore, the frontness of the stressed vowel /u/ (i.e., /u/ is less backed than /o/, but more backed than /ɔ/) was observed in this study for both genders, and tend to be more pronounced in older males (where /u/ is less backed than the other two back vowels). This was also observed in some inquiry points of the *Alto-Minho and Trás-os-Montes* and in the Portuguese central-southern dialects (Brissos, 2020; Brissos & Saramago, 2014). The palatalization of the stressed vowel /u/ was verified previously in the interior-center, in the *Barlavento Algarvio* and in the islands of Madeira and São Miguel (Azores) (Blayer, 2002; Brissos, 2011; Cintra, 1971; Cuesta & Luz, 1980; Segura, 2013). Nonetheless, to a lesser or greater degree, the palatalization of the vowel /u/ has also been reported in several other areas of the country, from north to south, especially in Coimbra region (Boléo & Silva, 1961; Cuesta & Luz, 1980; I. A. Santos, 2003).

Another issue, related with phonetic changes, is that the central high vowel [ɨ] was produced with a particularly low F1 value, and the low vowel usually transcribed as [ɐ] presents rather the values of a central vowel [ə], regardless of age. Pellegrini et al. (2013) observed the same tendency, mainly in younger speakers. As in Pellegrini et al. (2013), this acoustic correlate supports the idea in favor of the [ɨ] symbol (Veloso, 2007), whereas the central position occupied by the vowel [ɐ] suggests that the [ə] symbol might be more appropriate. Brissos (2018) also indicated that, in the Portuguese tradition, the symbol [ɐ] is used to represent a more closed vowel than a direct reading of the IPA supposes.

7.3.4 Study limitations

Given the methodological differences across previous studies, variable results are not surprising. For that, it is difficult to fix a particular age or age range where changes occur in both genders (Kent &

Vorperian, 2018).

Speaker age leaves traces in all phonetic dimensions and its impact on the speech is influenced by numerous factors, such as physiological condition, occupations and lifestyle habits (Makiyama & Hirano, 2017; Schötz, 2006), which were not handled in this study. Also the type of speech samples used could affect the results. In more conversational contexts, speakers tend to show decreases in average vowel duration coupled with a higher degree of vowel centralization (Fletcher et al., 2015). It is possible that, in order to see differences in vowel centralization with age, a task which demands greater movement of speakers' vocal tracts might be required (Fletcher et al., 2015). And finally, vowel duration was not controlled, which renders comparisons across studies to be problematic in several ways.

AGE AND VOWEL CLASSIFICATION IMPROVEMENT BY THE INCLUSION OF VOWEL DYNAMIC FEATURES

The main body of this chapter is based on the study published in International Journal of Speech Technology[1].

Despite speech being inherently dynamic (Yuan, 2013), most of the acoustic studies and age classification experiments have focused on static features to characterize vowel acoustics. Thus, the static approach to extract vowel formants has dominated the acoustical vowel studies, but it is a convenient simplification (Almurashi et al., 2019; C. I. Watson & Harrington, 1999), since vowel formants are not static but display vowel inherent spectral change (VISC) over time (Williams & Escudero, 2014) (the time-varying nature of vowel formants 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").

Additionally, as in Eichhorn et al. (2018), the results presented in the previous Chapter (Chapter 7), also indicated that older speakers might present specific articulatory adjustments during speech. Taking into account the aforementioned, a dynamic approach to study the aging effect on vowel production could be important (Gahl & Baayen, 2019), but the effect of aging on dynamic properties of vowels has not been extensively studied.

The present study aims to explore the usefulness of dynamic features for classification tasks and characterization of vowel acoustics. This study extends the previous results (Chapter 7) by using a different methodology that takes into account not only the static cues, but also formant dynamics of EP vowels to study the age effects, as differences between vowels produced by different age groups could potentially be manifested both in the general location of the vowels in the vowel space, as well as in the direction and extent of the formant movement.

Specifically, this study intends to:

- explore the effect of static and dynamic features on the classification of age and vowels;
- analyze the percentage of errors in vowel classification by age group and type of features (static or dynamic);
- statistically study the effects of age and gender on dynamic properties of formant frequencies (F1, F2 and F3).

[1] Albuquerque, L., Teixeira, A., Oliveira, C., & Figueiredo, D. (2022). Age and vowel classification improvement by the inclusion of vowel dynamic features. *International Journal of Speech Technology*, 25(1025-1040).

8.1 Dynamic extraction of vowel formant frequencies

8.1.1 Parameters for the characterization of dynamics

Two sets of approaches have been developed to extract the vowel dynamic information: a series of successive time points (multiple time point approach) and by curve-fitting (Van der Harst & Van de Velde, 2014).

In the multiple time point approach, measures may involve comparing formant frequencies from two discrete time points (at the beginning and at the end of a vowel token) (Morrison & Assmann, 2013), or from more than two time points (e.g., 5 points (Fox & Jacewicz, 2009); 30 points (Williams & Escudero, 2014)). For example, in the three-point model the formant measures are taken from three locations, namely, at 20% onset, 50% midpoint, and 80% offset during vowel duration (Almurashi et al., 2019; Hillenbrand, Getty, Clark, & Wheeler, 1995). Beyond that, dynamic variations in the formants F1 and F2 may be characterized by different measures such as trajectory length (TL) and the spectral rate of change (spectral roc) (Fox & Jacewicz, 2009; Williams & Escudero, 2014). Although these latter metrics incorporate more detailed spectral information, they do not account for the directionality of the change (Williams & Escudero, 2014). In other words, they do not account if the frequencies actually increase or decrease over time (Williams & Escudero, 2014).

Furthermore, vowel dynamics can also be expressed by curve-fitting parameterizations, by fitting parametric curves such as polynomials (Themistocleous, 2017) or discrete cosine transforms (DCT) (Elvin, Williams, & Escudero, 2016; Sarvasy, Elvin, Li, & Escudero, 2020; C. I. Watson & Harrington, 1999; Williams & Escudero, 2014) to formant contours for describing the shape of complex curve (Van der Harst & Van de Velde, 2014). These approaches allow to analyze formant means as well as the direction and magnitude of the formant change (Ewald, Liina Asu, & Schötz, 2017; Williams & Escudero, 2014). Also, these approaches avoid an arbitrary choice of the vowel targets, which can be problematic when a vowel appears not to have a steady-state section, or the formants reach a minima or maxima at different times (C. I. Watson & Harrington, 1999).

Approaches only with two time points have performed equally well as more sophisticated parametric curve approaches in distinguishing vowel categories (Morrison, 2013; Zahorian & Jagharghi, 1993). However, a whole trajectory approach based on parametric curves outperforms a two time point approach to extract speaker information (Morrison, 2013). Thus, the measurements at more time points pay off in return to more information about the speaker (such as regional origin) than static approaches (Van der Harst & Van de Velde, 2014).

In the current study, formant trajectories will be modelled as DCT coefficients, which is a method that has previously been used for the analysis of vowels of other languages (e.g., Elvin et al. (2016); Ewald et al. (2017); Sarvasy et al. (2020); Williams and Escudero (2014)). The DCT, a method proposed by Harrington, Cassidy, and Cassidy (1999), decomposes the signal into a set of cosine waves, whereby the resulting amplitudes of these cosine waves are the DCT coefficients (Jannedy & Weirich, 2017). The first three DCT coefficients can be interpreted as: the formant trajectory mean (C0); the magnitude and direction of change from the mean (i.e., the trajectory slope) (C1); and the curvature of the formant trajectory (C2) (Elvin et al., 2016; Sarvasy et al., 2020; Williams & Escudero, 2014). DCT coefficient values above the C2 represent the amplitudes of higher frequency cosines and therefore correspond to complex aspects of the shape of the formant trajectories (Elvin et al., 2016; Jannedy & Weirich, 2017; Morrison, 2013; Sarvasy et al., 2020; Williams & Escudero, 2014). DCTs is the best procedure for quantifying the shape of complex curve (Morrison, 2013; Van der Harst & Van de Velde, 2014), since many of the other approaches have theoretical or practical flaws (Morrison, 2013). Morrison (2013) indicates that there is a conceptual problem with the combination of two canonical steady-state targets and a linear fit to the whole trajectory (e.g., approaches with two time points).

8.1.2 Dynamics and vocoids classification

Although static vowel features may be sufficient for vowels classification, as it has previously been shown for some languages (Almurashi et al., 2019; Sarvasy et al., 2020; Williams, Van Leussen, & Escudero, 2015), a number of studies have reported that vowel dynamic cues contain essential information, not only for diphthongs but also for monophthongs (Almurashi et al., 2019). Furthermore, formant dynamics have been useful for: enhancing the classification of vowels based on acoustics for some languages (Almurashi et al., 2019; Al-Tamimi, 2007; Chittaragi & Koolagudi, 2019; Elvin et al., 2016; Jacewicz & Fox, 2012, 2013; Williams & Escudero, 2014; Yuan, 2013; Zahorian & Jagharghi, 1993)); determining cross-dialectal acoustic differences for some vowels (Jacewicz & Fox, 2013; Van der Harst & Van de Velde, 2014; Williams & Escudero, 2014); improving the within-class separation of the Australian English tense vowels from their lax counterparts (C. I. Watson & Harrington, 1999); and providing vowel phonetic details that are relevant in second-language (L2) learning (S.-H. Jin & Liu, 2013). However, there is no acoustic evidence, due to a lack of research in this field, that formant dynamics may actually vary systematically across age groups and that the use of time-varying features may be a subject to age variation (Albuquerque, Oliveira, Teixeira, & Figueiredo, 2021).

Dynamic acoustic features for vowel classification has not presented the same importance for all languages (Sarvasy et al., 2020), and typically, the first and second DCT coefficients (i.e., C0 and C1) are sufficient for classifying vowels (Elvin et al., 2016; Williams & Escudero, 2014), whereas formant curvature (C2) may not be necessary (Elvin et al., 2016; Williams & Escudero, 2014). For example, in Australian English, formant trajectory means, duration, magnitude and direction of formant trajectory slope are essential acoustic parameters for representing vowels (Elvin et al., 2016). Nonetheless, C2 can aid with more fine-grained/ subtle phonetic information specific of the speakers (e.g., dialect, age or gender information) (Elvin et al., 2016).

8.2 Method

In this section, the speakers' profile and speech database are given in *dataset*. Then, the parameters extraction method and the features sets used is explained in the section of acoustic analysis and features. Finally, the classification procedure section includes the classifiers and the model evaluation.

8.2.1 Dataset and age groups division

The EP speech database used in this study is based on the segmental data produced by the 112 native Portuguese speakers, and have previously been described in Chapters 4 and 5.

Briefly, the speech corpus consisted of 36 disyllabic words, with the EP vowels [i], [e], [ɛ], [a], [o], [ɔ] and [u] in stressed position and the vowels [ɨ] and [ɐ] in unstressed position, in stop and fricative consonant context. Each participant produced 12 repetitions of each vowel, in a total of 108 productions by speaker (112 participants x 36 words x 3 repetitions = 12 096 recordings).

As the understanding of how the vowel formant dynamics changes with age is the main purpose of this work, but only a relatively small amount of data is available for each age, it is necessary to define age groups. Four different age divisions were defined, with the objective of analyzing which age division fits better the lifespan vowel changes. Table 8.1 summarizes the allocation of age ranges to different age groups. Note that, on average, each speaker has 100 vowel productions, i.e., to obtain the approximate value of data used in each experiment, it is necessary to multiply the number of speakers by 100.

Table 8.1: Description of the age-groups divisions used in the experiments and gender distribution. N corresponds to the number of groups in the experience. (adapted from Albuquerque, Teixeira, et al. (2022))

N	Designation	Age Groups			
2	**Non older/ Older**	**[35-64]**	**[65-97]**		
		30 female; 30 male	26 female; 26 male		
2	**Younger/ Older**	**[35-45]**	**[70-97]**		
		11 female; 9 male	21 female; 18 male		
3	**Younger/ Middle-aged/ Older**	**[35-45]**	**[46-69]**	**[70-97]**	
		11 female; 9 male	24 female; 29 male	21 female; 18 male	
4	**Younger/ Middle-aged/ Older/ Oldest-old**	**[35-49]**	**[50-64]**	**[65-79]**	**[80-97]**
		15 female; 15 male	15 female; 15 male	16 female; 15 male	10 female; 11 male

From now on, in order to simplify, the four age divisions were designated based on the number of age groups compared, i.e., 2 groups *(non older/ older)*, 2 groups *(younger/ older)*, 3 groups and 4 groups.

Briefly, the main assumptions for these age divisions rely on the chronological definition of older people, and also on the reported patterns of physiological changes along the aging continuum (He et al., 2016; Mautner, 2011; World Health Organization, 2007). The experience with 2 groups *(non older/ older)* was defined based on the World Health Organization definition of older persons (cf. Section 2.1.1) (He et al., 2016; United Nations, Department of Economic and Social Affairs, Population Division, 2019). To allow a comparison between individuals at different stages in the aging process, and due to the large age-range in the classification on *non older/ older*, an experience with 4 groups was performed. The youngest group [35-49] represents adults prior to the onset of physical aging that involve changes on the vocal fold tissue (Mautner, 2011). The oldest group [80-97] represents healthy adults who are at an age where changes on the vocal fold tissues tend to be more prevalent in the literature (Mautner, 2011).

Additionally, voice is influenced by hormones and women experience a menopause around the age of 51 (with 3 to 4 years interval deviation) (Lã & Ardura, 2020). Therefore, the experience with 3 groups was designed to take into account this issue. Based on the same assumption and to test age groups with more expected distinctive characteristics, an experience with 2 groups *(younger/ older)* was designed removing the speakers of the middle-aged [46-69].

8.2.2 Acoustic analysis and features

The median values of F1, F2 and F3 were extracted from the central 40% of each target vowel using a procedure adapted from Escudero et al. (2009). This procedure automatically define the formant ceiling of the LPC analysis based on within-speaker and within-vowel variation, as described in Section 5.3.1 (Segmental Parameters). Thus, for each vowel produced by each speaker there is only one "optimal ceiling". The "optimal ceiling" was used to extract the vowel formant frequencies at 35 equally spaced time points within the central 60% of each token (adapted from Williams and Escudero (2014)), in order to reduce the effects of possible formant transitions associated with flanking consonants.

In order to characterize F1, F2, and F3 trajectories, each set of 35 formant values from each vowel token were fitted using DCT with the R package `emuR` (Winkelmann, Jaensch, Cassidy, & Harrington, 2020). Each DCT coefficient obtained, which characterizes an aspect of a formant trajectory's shape based on a cosine (Sarvasy et al., 2020; Williams & Escudero, 2014), was included as input parameters on classification models.

Specifically, the C0, C1, and C2 coefficients were used in this study to analyze the age effect on EP vowel formant trajectories, since for other languages, DCT coefficients below the C2 are typically sufficient for characterizing formant trajectory shape in vowels (Elvin et al., 2016; Williams & Escudero, 2014). Different combinations of the features (duration and the three DCT coefficients (C0, C1 and C2) of vowel formants (F1, F2 and F3)) were employed as predictors. The feature sets used in this study are described in Table 8.2.

Table 8.2: Description of feature sets used as predictors. (adapted from Albuquerque, Teixeira, et al. (2022))

Feature sets	**Definition**
Baseline 1 (B1)	C0 of F1 and F2 *(which corresponds to static information of F1 and F2)*
B1C1	Baseline 1 + C1 of F1 and F2
B1C1C2	Baseline 1 + C1 and C2 of F1 and F2
B1_d	Baseline 1 + vowel duration
B1C1_d	Baseline 1 + C1 of F1 and F2 + vowel duration
B1C1C2_d	Baseline 1 + C1 and C2 of F1 and F2 + vowel duration
Baseline 2 (B2)	C0 of F1, F2 and F3 *(which corresponds to static information of F1, F2 and F3)*
B2C1	Baseline 2 + C1 of F1, F2 and F3
B2C1C2	Baseline 2 + C1 and C2 of F1, F2 and F3
B2_d	Baseline 2 + vowel duration
B2C1_d	Baseline 2 + C1 of F1, F2 and F3 + vowel duration
B2C1C2_d	Baseline 2 + C1 and C2 of F1, F2 and F3 + vowel duration
C1 (F1-F2)	C1 of vowel formants (F1 and F2) *(which corresponds to the trajectory slope)*
C1 (F1-F3)	C1 of vowel formants (F1, F2 and F3) *(which corresponds to the trajectory slope)*
C2 (F1-F2)	C2 of vowel formants (F1 and F2) *(which corresponds to the curvature of the formant trajectory)*
C2 (F1-F3)	C2 of vowel formants (F1, F2 and F3) *(which corresponds to the curvature of the formant trajectory)*
Duration (d)	Vowel duration (ms)

8.2.3 Classification procedure

The existence of a significant age effect in vowels' dynamics was indirectly studied by measuring the impact of using static and dynamic information in two classification tasks: age classification and vowel classification.

The classification tasks by each gender were conducted with the following classifiers: Discriminant Analysis, Decision Tree Model, Naïve Bayes, Support Vector Machine and Ensemble classifiers. The 17 features sets described above were employed as predictors.

Set of Classifiers

In a first step, an initial experiment with 2 age groups ([35-64] vs [65-97]) with the Classification learner app in Matlab (version R2018b) was performed using 5-fold cross-validation, and the sub-models that presented the best accuracy were selected for the study.

These classifiers were used in Matlab scripts to perform all the classification tasks. That is, in order to determine the potential usefulness of various DCT coefficients and duration as measures for carrying out age classification, the selected classification algorithms were run using all the defined feature sets to test how accurately combinations of these input variables classify tokens by age group.

Additionally, the selected classification algorithms were run using all the defined feature sets on vowel classification task, to examine the vowel classification performances obtained, and the percentage of vowel classification errors were examined by age group for each type of information (i.e., static or dynamic features).

The selected five classifiers are briefly described below:

Quadratic Discriminant Analysis (QDA) – QDA is one type of discriminant analysis which creates non-linear boundaries between classes of samples (elipse, parabola or hyperbola) (The MathWorks, 2021). That is, boundary divides the space into regions, with each according to different classes. The QDA model assumes that different classes generate data based on different Gaussian mixture distribution (gmdistribution). Also QDA assumes different variance–covariance matrices for each class (The MathWorks, 2021). The Matlab function used was `fitcdiscr(features, classes, 'DiscrimType', 'quadratic')`.

Fine Gaussian Support Vector Machines (fgSVM) – SVM is a classification method that separates the classes by constructing a hyperplane (The MathWorks, 2021), which works as discriminative classifier (Dey & Alam, 2018). The training of the model seeks to find the hyperplane or group of hyperplanes that separate the classes with the widest margin. The Matlab functions `templateSVM` and `fitcecoc` were used. Also the Matlab version designated as "Fine Gaussian SVM" was selected due to its higher flexibility, as it makes finely detailed distinctions between classes, with the Gaussian kernel function and the kernel scale set to $\sqrt{P}/4$, where P corresponds to the number of predictors used (The MathWorks, 2021).

Kernel Naïve Bayes (KNB) – KNB is one type of Naïve Bayes classifier, so called because it is a Bayesian classifier that makes a simplifying (naïve) assumption about how the features interact. This algorithm leverages Bayes theorem and makes the assumption that predictors are conditionally independent, given the class (The MathWorks, 2021). In KNB classifier the kernel smoother type is Gaussian and the kernel smoothing density support is Unbounded. The Matlab function used was `fitcnb`.

Fine Tree (FT) – FT fits binary decision tree for multiclass classification. The classification trees are trained to predict responses to data. To predict a response, follow the decisions in the tree from the root (beginning) node down to a leaf node. The leaf node contains the response. Each step in a prediction involves checking the value of one predictor (The MathWorks, 2021). The Matlab function used was `fitctree`, and the Matlab version designated as ”Fine Tree” was selected due to its higher flexibility, as it uses many more leaves to make many fine distinctions between classes (maximum number of splits is 100) (The MathWorks, 2021). The Gini’s diversity index (gdi) is the split criterion measure used for deciding when to split nodes.

Ensemble Bagged Trees (EBT) – The Ensemble classifiers combine the results from different learners producing a better accuracy (The MathWorks, 2021). The EBT method uses Breiman’s ’random forest’ algorithm (Breiman, 2001). The Matlab functions `templateTree` and `fitcensemble` were used.

Model evaluation

The five classification algorithms were employed for the four different age classification tasks, and also for vowel classification, by gender. All the defined feature sets were employed as predictors.

In order to evaluate the classification results from all classifiers, the data were separated into a training set consisting of 90% of the data and into a test set consisting of 10% of the data (the number of productions for training and testing varies with the type of experience (i.e., with the number of speakers involved). Also, 10-fold cross validation was employed. All experiments were performed in Matlab (version R2019b) using default configurations.

The classifier is trained using the training set and the classification performance is estimated on the test set. After classification, confusion matrices have been computed. For model comparison the accuracy, that is the proportion of total number of corrected predictions, was obtained.

Due to the different numbers of classes between experiences, the performance of the classifier is also assessed by how strongly correct classification rate departs from the chance-level rate achieved by a classifier that would randomly associate the samples to the various classes. For a given number of classes c in the experience, the percent theoretical chance level of classification is given by $100/c$. The chance-level accuracy was computed as the difference between the accuracy reported and the theoretical percent chance-level associated to each classification task.

Additionally, in vowel classification, the prediction error (i.e., the mean of misclassified samples) by age group was analyzed.

As secondary measure, the mean CPU time to perform each feature set by classification task was extracted, but not analyzed in detail. The classification tasks were performed using an Intel Xeon E3-1231 and an Intel i7-7500U, with Matlab 2019b.

Table 8.3: Performance of all classifiers for the five classification tasks. Note that the chance-level accuracy was computed based on the mean accuracy.(adapted from Albuquerque, Teixeira, et al. (2022))

		2 age groups (non older/older)		**2 age groups** (younger/older)		**3 age groups**		**4 age groups**		**9 vowels**	
Classifiers	**Metrics**	**Female**	**Male**	**Female**	**Male**	**Female**	**Male**	**Female**	**Male**	**Female**	**Male**
QDA	**Mean accuracy**	58%	57%	64%	64%	45%	52%	33%	31%	72%	72%
	Chance-level accuracy	8%	7%	14%	14%	12%	19%	8%	6%	61%	61%
	Maximum accuracy	63%	61%	68%	68%	48%	57%	38%	34%	91%	91%
fgSVM	**Mean accuracy**	**63%**	**62%**	**71%**	**71%**	**51%**	**57%**	**40%**	**39%**	71%	70%
	Chance-level accuracy	**13%**	**12%**	**21%**	**21%**	**18%**	**24%**	**15%**	**14%**	60%	59%
	Maximum accuracy	71%	70%	79%	76%	58%	62%	51%	50%	91%	90%
KNB	**Mean accuracy**	59%	58%	66%	66%	46%	54%	35%	32%	**73%**	**72%**
	Chance-level accuracy	9%	8%	16%	16%	13%	21%	10%	7%	**62%**	**61%**
	Maximum accuracy	64%	61%	69%	68%	48%	57%	39%	35%	90%	90%
FT	**Mean accuracy**	59%	58%	66%	66%	46%	53%	35%	33%	72%	71%
	Chance-level accuracy	9%	8%	16%	16%	13%	20%	10%	8%	61%	60%
	Maximum accuracy	64%	62%	72%	70%	50%	57%	40%	37%	88%	88%
EBT	**Mean accuracy**	61%	60%	68%	69%	48%	54%	38%	37%	72%	72%
	Chance-level accuracy	11%	10%	18%	19%	15%	21%	13%	12%	61%	61%
	Maximum accuracy	69%	68%	77%	78%	58%	63%	49%	48%	92%	91%

8.3 Results

Before analyzing the effects of static and dynamic properties of vowels on classification tasks, the general performance of the five classifiers is compared. The classifier that presents the highest accuracies was chosen for the study (i.e., for age classification tasks and vowel classification task). The age classification tasks were performed for all the feature sets and also for the four age divisions. The age division that obtained the best accuracy in age classification was used in subsequent experiments, namely to analyze the vowel classification errors by age, and to the statistical analysis of age and gender effect on vowel dynamic features.

8.3.1 General performance of the classifiers

In Table 8.3 the general performance of the classifiers is presented, namely the mean accuracy, the chance-level accuracy and the maximum accuracy obtained for each classification task (i.e., four different age classification tasks and a vowel classification task). Each classification task was performed by gender for all the 17 feature sets. The mean accuracy reported on Table 8.3 was computed as the average of the mean accuracy obtained for each feature set tested, per gender and classifier, for each classification task. The chance-level accuracy was computed based on the mean accuracy reported for each classification task. The maximum accuracy corresponds to the higher accuracy obtained for each classification task.

In general, the fgSVM classifier presents higher accuracies, namely in age classification tasks. For vowel classification, the results among classifiers are similar, but the KNB classifier presents the best accuracy, namely for female vowel classification.

The mean CPU time tend to vary between 0.2 s for QDA and 43 s for fgSVM, per each feature set classification. Despite the CPU time needed, the fgSVM classifier was chosen for all age classification tasks, and the KNB classifier was used in vowel classification study.

Additionally, all age classification tasks were also performed with the information of vowel type as predictor. However, this parameter did not have a significant effect on the classification results, so it was not used in this study.

8.3.2 Age classification: Effects of age and dynamic features

The effect of dynamic features in age classification (with the classifier selected in the previous section) for the four different age divisions is presented in Figure 8.1. Despite each age classification task being performed by gender for all the 17 feature sets, Figure 8.1 only presents results for some feature sets, as the main objectives are to analyze the general performance and also the accuracy improvement when adding dynamic information to the baselines (i.e., static features).

The results of mean accuracy (at left graphs of Figure 8.1) show a tendency for improvement when adding C1, and C1C2 to both baselines. Also duration contributes to improvements even when added after C1 and C2. The graph at right (Figure 8.1) make clearer the improvements by presenting the accuracy improvement (difference for the relative baseline). In general, the age classification for females presents a higher improvement when adding dynamic information (C1 and C2) than males.

The higher improvement obtained with the dynamic information added to the baselines B1 and B2 is observed for females in the experiment with 2 age groups *(younger /older)*. That is, when adding C1C2 to baseline 2, the improvement is approximately 9%. For males, the higher improvement is obtained in the experiment with 4 age groups, with an improvement of approximately 7% when adding C1C2 to baseline 1.

In general, the feature set that presents the highest accuracies for all the age classification experiments is B2C1C2_d for both genders, except for males on classification in three and four age groups. In these experiments, the feature set that obtain higher mean accuracy is B2C1_d.

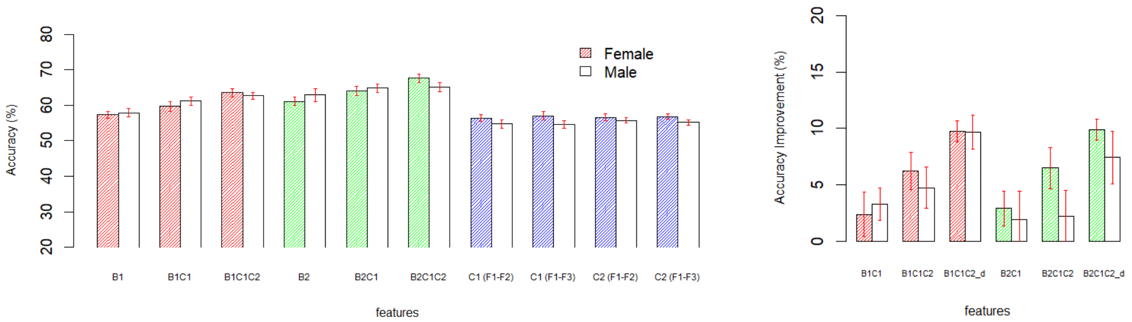
The age group that presents the highest mean accuracy on age classification is the division in *(younger/ older)*, with an accuracy of 79% and 76% for females and males, respectively (in the feature set B2C1C2_d). However, in the division of speakers in 3 age groups the accuracy improvement comparing to the chance level (i.e., the chance-level accuracy) is higher, namely for males, with a mean accuracy of 62% for the feature set B2C1_d. Females present a mean accuracy of 58% for the feature set B2C1C2_d.

8.3.3 Vowel classification: Effects of age and dynamic features

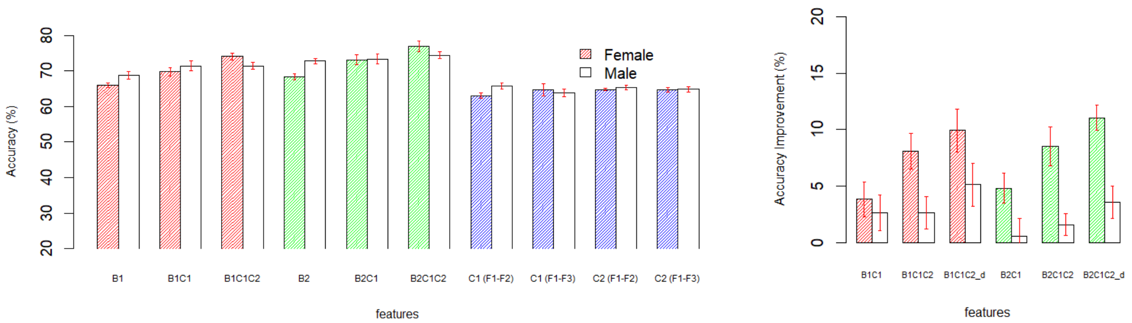
First, to explore the classification of the EP oral vowels according to their static and dynamic properties, the classification task was performed for all the speakers by gender, with the classifier KNB. As in age classification, despite vowel classification task being performed for all the 17 feature sets, the Figures 8.2 and 8.3 only present results for the more relevant feature sets.

Figure 8.2 shows that the mean accuracy of the vowel classification tends to slightly increase when adding dynamic information, namely the C1 coefficients, to the baselines B1 and B2. The accuracy improvement was around 2% for both genders. On the other hand, the addition of C1C2 to the baselines presented different tendencies between genders. For males, the accuracy improvement was approximately 3% in relation to the baselines. For females, when adding C1C2 to the baselines, the accuracy improvement is similar or lower than when adding only C1 to the baselines.

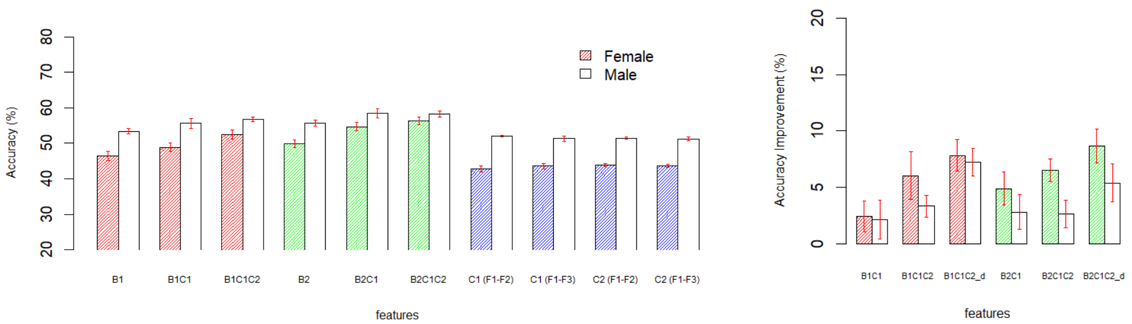
(a) 2 groups *(Non older/ Older)*



(b) 2 groups *(Younger/ Older)*



(c) 3 groups



(d) 4 groups

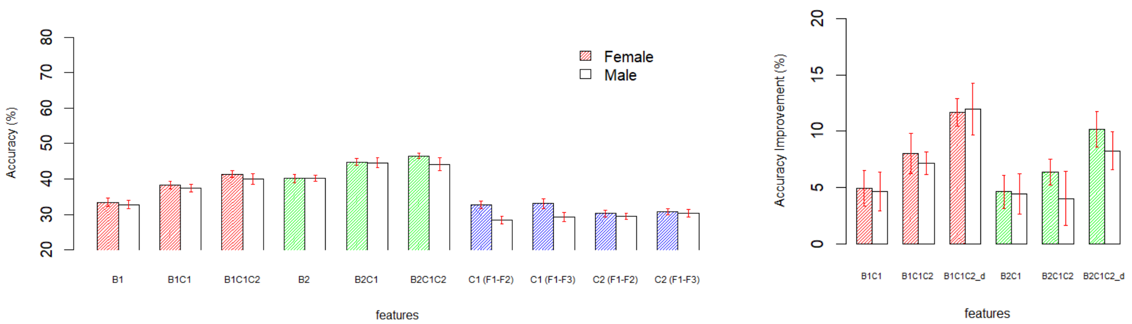


Figure 8.1: Age classification accuracy of fgSVM classifier for the 4 different age divisions considered. Bar graphs at left present the accuracy obtained for some feature sets; at right are presented the improvements regarding the correspondent baseline (differences from baselines). Male data is represented with white bars. Error bars in the graphs represent 95% CI. (adapted from Albuquerque, Teixeira, et al. (2022))

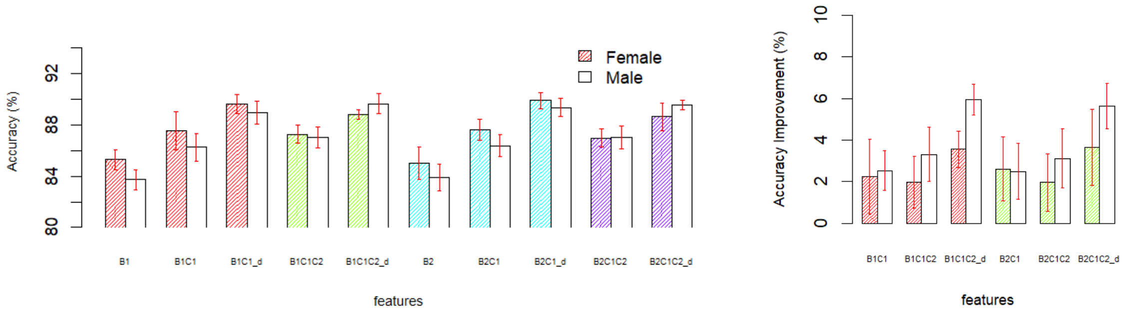


Figure 8.2: Vowel classification accuracy with kernel Naïve Bayes (KNB) classifier as function of features and gender. Bar graphs at left present the accuracy obtained for some feature sets; at right are presented the improvements regarding the correspondent baseline (differences from baselines). Male data is represented with white bars. Error bars in the graphs represent 95% CI. (adapted from Albuquerque, Teixeira, et al. (2022))

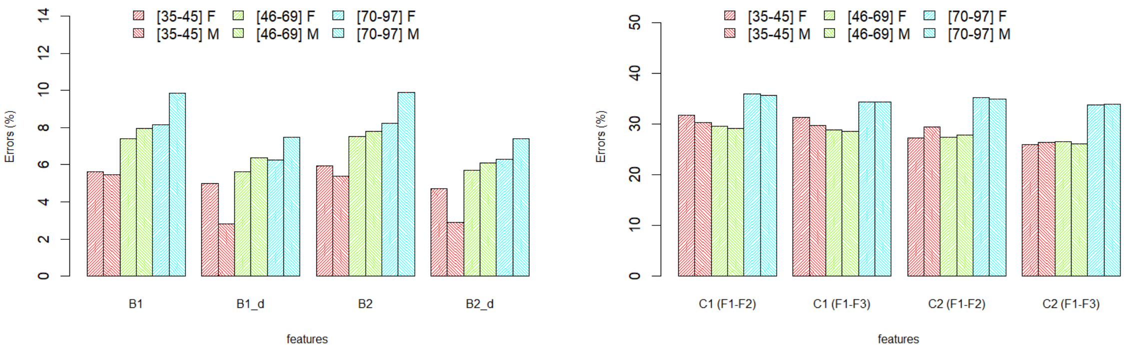


Figure 8.3: Vowel classification error with Kernel Naïve Bayes (KNB) classifier by age group and gender: At left, baseline (static) parameters (with and without vowel duration information); at right, dynamic features. (adapted from Albuquerque, Teixeira, et al. (2022))

The feature set that obtained the best mean accuracy for vowel classification was B1C1C2_d for males and B2C1_d for females. Nonetheless, the results of the feature set B2C1C2_d for males and B1C1_d for females show similar results, with all mean accuracies around 90%. Unlike F3, adding duration seems to improve the vowel classification accuracy.

More relevant than the overall accuracy is to analyze the differences in error classification by age, using dynamic information compared with static. For the vowel classification task, the classifiers have been trained on the whole speaker set, separately for each gender. The percentage of errors was analyzed based on the division of the speakers in 3 age groups (i.e., [35-45], [46;69], [70-97]), since, this age group division presented the best results on age classification for all speakers of the database.

The vowel classification errors by age and type of information (i.e., static parameters (B1 and B2) or dynamic features (C1 and C2 of vowel formants)) are presented in Figure 8.3. Even though, the vowel classification errors are higher when only dynamic information is used, it can be observed that variation of the percentage of errors with age depends on the type of information used (i.e., static or dynamic). The vowel classification errors based on static features tend to increase continuously with age, in both genders. Based on dynamic cues, the percentage of errors tend to be similar between the age groups [35-45] and [46-69], and only tend to increase more sharply in the last age group [70-97].

8.3.4 Age effect on dynamic cues

As a positive impact of using dynamic information in age classification tasks was observed, mainly for females, to complement the classification experiments the effects of age (and gender) on dynamic features was also investigated using an analysis of variance (ANOVA). The statistical analysis was conducted with the SPSS software package (SPSS 26.0). The mean of each dynamic feature (i.e., F1C1, F1C2, F2C1, F2C2, F3C1 and F3C2) was obtained for all vowels. The correlation between the dependent variables was evaluated by Spearman correlation. The correlations varied between −0.300 (F1C1 vs F1C2) to 0.273 (F1C1 vs F3C2), indicating low, non-significant correlation between them. For each dependent variable a two-way ANOVA was applied, with age and gender (between-subject factors). Due to the multiple testing (six models in total), the significant level used was 0.008 (i.e., 0.05/6). The ANOVA assumptions of residual normality and homogeneity of variance were verified. As in the previous section, the division of the speakers in 3 age groups was adopted.

The results of the ANOVAs are presented in Table 8.4. A main effect of age was revealed only for F1C1. Interestingly, the C2 of F1 and F2, but not the C1 of F1 and F2, show a main effect of gender.

Table 8.4: Results of the ANOVA by age, gender, and age * gender interaction on vowel formats dynamic coefficients. (adapted from Albuquerque, Teixeira, et al. (2022))

Variable	Age	Gender	Age*Gender
F1C1	$\boldsymbol{F(2,106) = 5.615,\ p = 0.005}$	$F(1,106) = 0.068,\ p = 0.795$	$F(2,106) = 1.242,\ p = 0.293$
F1C2	$F(2,106) = 1.874,\ p = 0.165$	$\boldsymbol{F(1,106) = 9.656,\ p = 0.002}$	$F(2,106) = 0.985,\ p = 0.377$
F2C1	$F(2,106) = 1.575,\ p = 0.212$	$F(1,106) = 3.301,\ p = 0.072$	$F(2,106) = 1.433,\ p = 0.243$
F2C2	$F(2,106) = 0.752,\ p = 0.474$	$\boldsymbol{F(1,106) = 10.593,\ p = 0.002}$	$F(2,106) = 3.326,\ p = 0.040$
F3C1	$F(2,106) = 0.448,\ p = 0.640$	$F(1,106) = 1.453,\ p = 0.231$	$F(2,106) = 0.951,\ p = 0.390$
F3C2	$F(2,106) = 1.145,\ p = 0.322$	$F(1,106) = 1.833,\ p = 0.179$	$F(2,106) = 1.180,\ p = 0.311$

Significant differences are highlighted with bold (p<0.008).

The mean values and the error bars of the dynamic DCT coefficients that showed significant differences (F1C1, F1C2 and F2C2) are presented in Figure 8.4.

As seen in Figure 8.4, the mean values of F1C1 are similar between genders (female: −7.6 ± 7.1; male: −6.8 ± 4.6), but present significant differences with age. Speakers from the group [35-45] display the highest mean F1C1 (−3.3 ± 4.7). Moreover, the older groups present similar values ([46-69]: −7.8 ± 5.7; [70-97]: −8.5 ± 6.1). Post hoc tests confirm as significant the difference between the younger group [35-45] and the older groups.

The F1C2 and F2C2 present significant differences between genders, and males present the highest values (F1C2 – female: -6.3 ± 1.9; male: -5.3 ± 1.4; F2C2 – female: -4.6 ± 3.4; male: -1.9 ± 3.7). That is, the curvature of the trajectories of F1 and F2 is more pronounced in females. Although not significant, the gender differences in F1C2 tend to decrease within age, so males and females of the last age group present similar values of F1C2 (Figure 8.4b).

Despite F2C2 also not presenting significant differences across genders with age, the curvature of the trajectory of F2 tends to differ between genders with age. That is, F2C2 present similar values between genders in the age group [35-45] and increase for males and decrease for females, with age (Figure 8.4c).

8.4 Discussion

This paper presents several experiments on the effects of dynamic features in age and vowel classification tasks based on acoustic features extracted from EP oral vowels, as well as explores how age affects both dynamic features and classification. The aim of these experiments is to contribute to increase the existing knowledge regarding effects of aging in speech and provide basis to better support for older adults by speech technologies. To this purpose, a number of classification experiments using vowels data from a large sample of healthy Portuguese adults were performed.

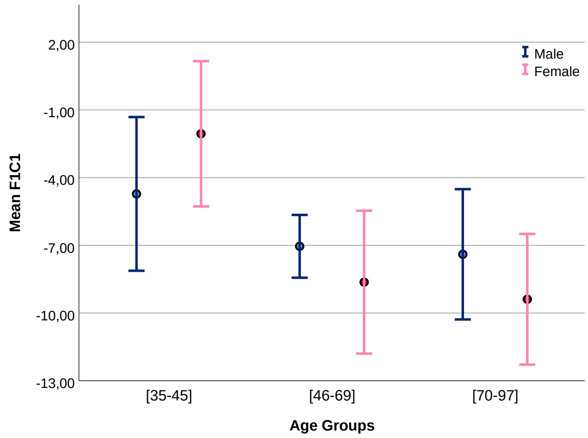
The study tested a representative set of classifiers (five) before focusing on just two of them: Fine Gaussian version of Support Vector Machines (fgSVM) for age classification and Kernel Naïve Bayes (KNB) for vowel classification, which were found to outperform the other classifiers for each classification task.

The classification tasks were attempted with different sets of features, starting with only static information for the first 2 or 3 formants and incrementally adding dynamic information provided by DCT coefficients C1 and C2. The inclusion of vowel duration was also investigated.

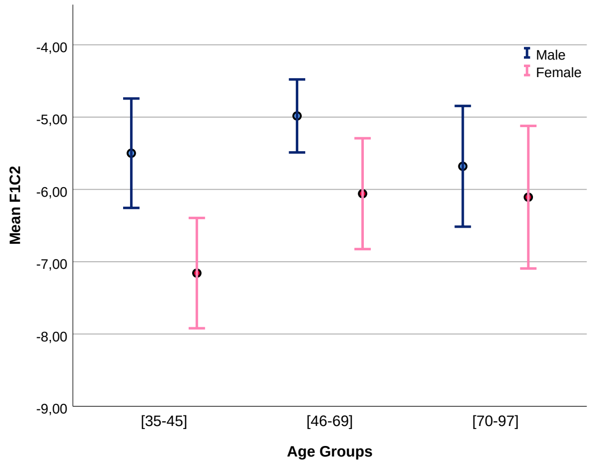
This study provides additional support on the relevance of vowel dynamics for age classification. Therefore, this constitutes a novel line of research for at least EP since most of the previous studies rely on static approaches to characterize the vowels (Albuquerque et al., 2014; Escudero et al., 2009; Oliveira, Cunha, et al., 2012; Pellegrini et al., 2013; Rato et al., 2017). The study shows the positive impact of adding dynamic cues of F1-F3 to static information in age and vowel classification tasks. Adding the dynamic features (C1 and C2) to static information resulted in higher accuracy improvement in age classification than in vowel classification, particularly for females.

The non-existence of consensus in age divisions motivated the adoption of several age groups for the classification experiments. The best results were obtained for the division in 2 groups *(younger/ older)*, with best accuracies, and 3 groups ([35-45], [46-69] and [70-97]), with better improvement relative to chance level, namely for males. These age divisions address the menopause stage for females: discarded in 2 groups division; only affecting one of the groups for the division into 3 groups. As expected, given that the middle-aged speakers have been excluded, the division in 2 groups *(younger/ older)* presented the higher accuracies. However, it is important to note that similar or higher improvements relative to chance level were obtained in the division in 3 groups, which indicated that this age division could fit better the lifespan speech alterations.

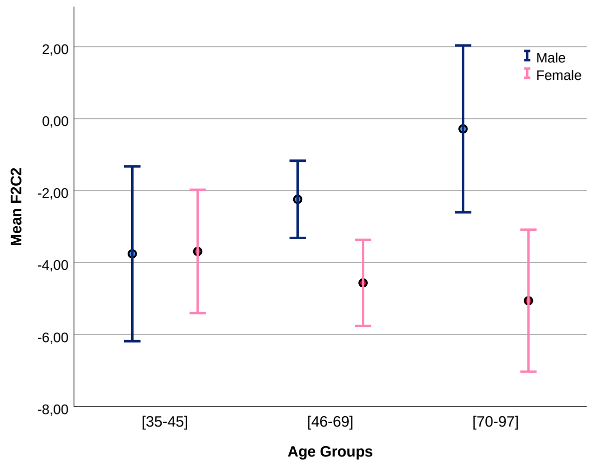
On the vowel classification task, similar mean accuracy was obtained for both genders (approximately 90%), with a tendency for information regarding F3 (both static and dynamic) not to improve the vowel classification performance. Despite the slight effects of dynamic information on vowel classification, the dynamic information tends to be more relevant for males. Also, the DCT coefficient C2, which represents the trajectory curvature, has a reduced influence on the vowel prediction, namely



(a) F1C1 by age and gender.



(b) F1C2 by age and gender.



(c) F2C2 by age and gender.

Figure 8.4: The effect of age and gender on mean values of F1C1, F1C2 and F2C2. The Error bars represent 95% CI. (adapted from Albuquerque, Teixeira, et al. (2022))

for females. As in Williams and Escudero (2014), C2 coefficients do not appear to contribute a great deal of extra information useful for characterizing all EP oral vowels.

The variation of vowel classification errors presented a different pattern between static and dynamic information with aging (3 age groups). While classification errors based on static features tend to increase continuously with age, based on dynamic cues, the percentage of errors only increases sharply in the last age group [70-97].

As expected, and reported for example in Almurashi et al. (2019), adding duration has a positive impact on vowel classification. The present work supports the extension of this effect to age classification, which being in accordance with previous studies for EP (Albuquerque et al., 2014; Pellegrini et al., 2013), and also for other languages that have shown a significant vowel duration increase with aging (Benjamin, 1982; D'Alessandro & Fougeron, 2018; Fletcher et al., 2015; Kyriaki, 2021; Mertens et al., 2020; Schötz, 2006; Tykalova et al., 2020).

To complement the classification experiments, ANOVA was performed on dynamic features with age and gender as factors, aiming to identify the dynamic features more affected by age. Only dynamic features related to F1 and F2 showed age or gender differences, providing additional support for the lower relevance of F3. As observed in Chapter 7, previous static studies have also reported that the frequencies of F3 were essentially unchanged with age (Eichhorn et al., 2018; Schötz, 2006). Interestingly, contrary to the trajectory slope (C1), the curvature (C2) of F1 and F2 presented gender differences, similarly to what has been reported for the static information of vowel formants (C0) (Eichhorn et al., 2018; Escudero et al., 2009). Furthermore, mean of F1C1, which is always negative, decreases with age, which means that the F1 trajectory slope increases with age for both genders. Jacewicz et al. (2011a, 2011b) have also observed that older speakers present greatest amount of spectral change, perhaps reflecting a somewhat "more exaggerated articulation" (Jacewicz et al., 2011a) in older speakers.

Globally, results tend to support the hypothesis that dynamic features of vowels carry important information about the speaker and are an interesting source of speaker-discriminating information (McDougall & Nolan, 2007).

Nonetheless, the current study has a number of limitations: the lack of balance in the age group divisions, the number of feature sets used, the type of speech task, and the non-control of some characteristics of the speech stimuli (e.g., word frequency, vowel duration (Gahl & Baayen, 2019; Munson & Solomon, 2004; Reubold & Harrington, 2015)). These factors could be problematic and affect the vowel formant frequencies variation.

This page intentionally contains only this sentence.

AGE-RELATED CHANGES IN EUROPEAN PORTUGUESE SEMI-SPONTANEOUS SPEECH

The main body of this chapter is based on the results of the study published in the journal LOQUENS [1].

Considering the paucity of literature and the conflicting findings reported in prior research, this study aims to contribute to increase knowledge regarding age effects on spontaneous speech of EP adults. That is, the objective of this study was to analyze the effects of healthy aging namely on the speech rhythm (duration measures, speech and articulation rate), intonation (Speaking F0) and voice quality (HNR) in semi-spontaneous speech, which could contribute to our understanding of the aging effects on natural spoken language production. Therefore, this study extends the research of Chapter 7 about the age-related effects on segmental level (i.e., vowel acoustic features in reading sentences).

9.1 Method

9.1.1 Speakers and speech data

The EP speech database used in this study have previously been described in Chapters 4 and 5. Briefly, the "Cookie Theft" picture description was performed by 112 healthy native Portuguese speakers, without voice or speech problems, aged between 35 and 97 (56 men and 56 women). For more details about the study design and recording protocol see Chapter 4, regarding segmentation and speakers' characterization see Chapter 5.

In order to analyze the age effects in rhythm, intonation and voice quality, several parameters (Table 9.1) were extracted from semi-spontaneous speech, based on the suprasegmental annotation of speech, pauses and syllables.

9.1.2 Statistical analysis

The statistical analysis was conducted with the SPSS software package (SPSS 25.0). For each suprasegmental parameter a multiple linear regression was conducted with the following explanatory variables: age (continuous), gender (male: reference group, female), and the interaction between age and gender. The regression coefficients and the correspondent 95%CI were calculated. The level of significance was $p < 0.05$ for all statistical analysis.

[1] Albuquerque, L., Valente, A. R. S., Teixeira, A., Oliveira, C., & Figueiredo, D. (2021). Age and gender effects in European Portuguese spontaneous speech. *LOQUENS*, 8(1-2), e077.

Table 9.1: Description of the suprasegmental parameters extracted. (adapted from Albuquerque, Valente, Teixeira, Oliveira, and Figueiredo (2021))

Parameters	Description
Total speech duration (s)	Sum of speech duration of all speech intervals
Total pause duration (s)	Sum of pause duration of all pause intervals
Total recording duration (s)	Sum of all speech and pause intervals
Percent pause time (%)	Total pause duration divided by total time (all speech and pause intervals)
Speech pause ratio	Total time talking divided by total pause time
Number of pauses	Number of pause intervals in the description task
Mean pause duration (s)	Duration average of pause length
Mean speech duration (s)	Duration average of speech length
Pause variability (s)	Standard deviation (SD) of pause length
Speech variability (s)	Standard deviation (SD) of speech length
Number of syllables	Sum of all syllable onset detected within all speech intervals
Speech rate (syllables/s)	Number of syllables divided by the total time (include pause intervals)
Articulation rate (syllables/s)	Number of syllables per seconds of speech without pauses
Speaking F0 (Hz)	Average number of vibrations per second of the vocal folds in the entire speech sample
HNR (dB)	Average ratio of the aperiodic energy to the harmonic energy

9.2 Results

This section presents the results of the selected acoustic parameters, and their statistical analyses. The multiple linear regression coefficients are displayed in Table 9.2 and the results revealed significant differences between genders for the parameters speech rate, articulation rate, speaking F0 and HNR. There was a significant age effect in males for total speech duration, percent pause time, mean pause duration, speech variability and number of syllables; and in females for the suprasegmental parameters: speech rate, articulation rate and HNR.

Table 9.2: Results of multiple linear regression: the effect of gender and gender*age interaction on suprasegmental parameters. Grey cells represent significant results ($p<0.05$). (adapted from Albuquerque, Valente, Teixeira, Oliveira, and Figueiredo (2021))

	Intercept			Gender (Male)			Male * Age			Female * Age		
	B	95% CI		B	95% CI		B	95% CI		B	95% CI	
Total speech duration (s)	29.20	15.58	42.82	7.24	-12.32	26.81	-0.25	-0.47	-0.03	-0.12	-0.33	0.10
Total pause duration (s)	6.93	0.88	12.99	-0.14	-8.84	8.56	0.03	-0.07	0.12	0.02	-0.08	0.12
Total recording duration (s)	36.13	18.57	53.70	7.10	-18.13	32.33	-0.22	-0.50	0.06	-0.10	-0.38	0.18
Percent pause time (%)	22.81	10.41	35.22	-10.08	-27.90	7.73	0.27	0.07	0.47	0.07	-0.13	0.26
Speech pause ratio	13.15	-2.37	28.67	-8.24	-30.54	14.05	-0.03	-0.28	0.22	-0.11	-0.35	0.14
Number of pauses	9.08	3.23	14.93	3.55	-4.85	11.95	-0.06	-0.15	0.03	-0.02	-0.11	0.08
Mean pause duration (s)	0.91	0.36	1.46	-0.53	-1.32	0.26	0.01	0.00	0.02	0.00	-0.01	0.01
Mean speech duration (s)	4.08	2.13	6.02	-1.26	-4.06	1.53	-0.01	-0.04	0.02	-0.02	-0.05	0.01
Pause variability (SD)	0.31	-0.30	0.92	-0.16	-1.04	0.72	0.01	0.00	0.02	0.01	0.00	0.02
Speech variability (SD)	1.71	1.08	2.34	0.35	-0.56	1.25	-0.01	-0.02	0.00	0.00	-0.01	0.01
Number of syllables	112.14	42.38	181.89	65.89	-34.30	166.08	-1.14	-2.25	-0.02	-0.14	-1.24	0.95
Speech rate (syllables/s)	2.40	1.46	3.34	2.03	0.68	3.39	-0.01	-0.03	0.00	0.02	0.00	0.03
Articulation rate (syllables/s)	3.06	2.04	4.08	1.97	0.50	3.44	0.00	-0.01	0.02	0.03	0.01	0.04
Speaking F0 (Hz)	204.29	181.36	227.22	-82.13	-115.06	-49.20	0.07	-0.29	0.44	-0.23	-0.59	0.13
HNR (dB)	17.56	15.29	19.82	-7.07	-10.32	-3.82	-0.00	-0.04	0.04	-0.05	-0.09	-0.02

B = Linear Coefficient

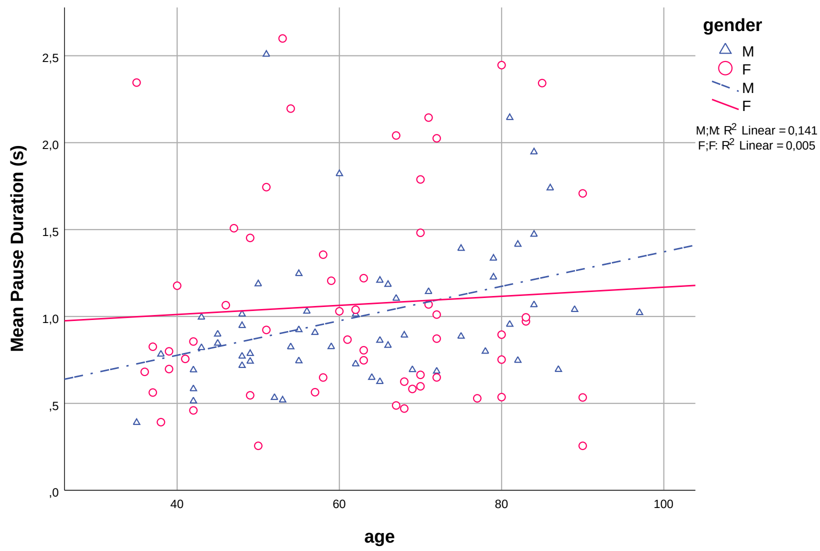


Figure 9.1: Scatterplot and regression lines for mean pause duration by age and gender. Each symbol corresponds to one speaker. Solid line and circles: females; dashed line and triangles: males. (adapted from Albuquerque, Valente, Teixeira, Oliveira, and Figueiredo (2021))

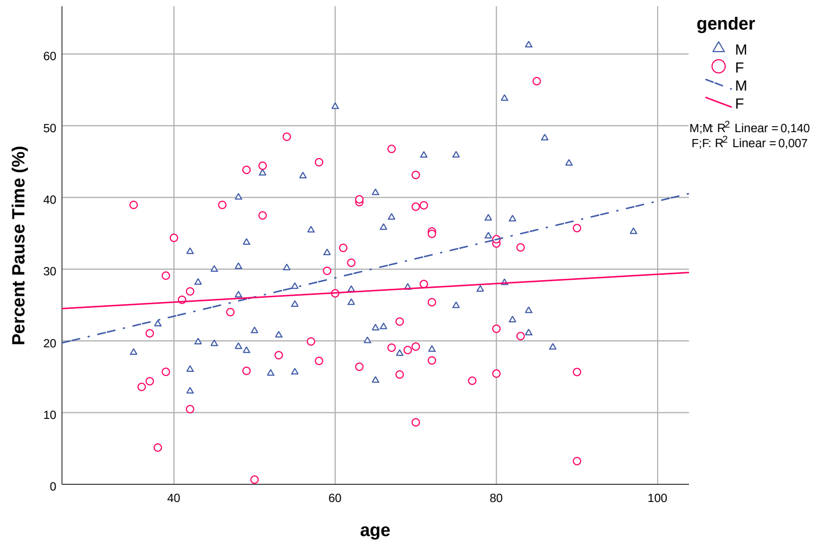


Figure 9.2: Scatterplot and regression lines for percent pause time by age and gender. Each symbol corresponds to one speaker. Solid line and circles: females; dashed line and triangles: males. (adapted from Albuquerque, Valente, Teixeira, Oliveira, and Figueiredo (2021))

9.2.1 Rhythmic parameters

Concerning the mean pause duration (see Figure 9.1) and the percent pause time (see Figure 9.2), the regression lines showed an increase with age, mainly in males, which means pause duration increases from approximately 0.7 s at age 35 to 1.4 s at age 100, and the percent pause time increases 20% in the same age interval. The multiple linear regression revealed that, for both parameters, the age effect is only significant for men.

Regarding articulation rate (Fig. 9.3): (1) age effect is gender dependent, with men presenting a

9.4% higher mean articulation rate than women; (2) difference between genders decreases with age, with older men and women presenting similar mean articulation rate, due to an increase in women's articulation rate. The regression model confirmed the gender differences and also showed a significant increase with age for female. Speech rate (Figure 9.4), as articulation rate, presented a tendency to increase in women. In men, a higher tendency to decrease, due to a rise in the percent pause time with age was observed.

Additionally, a decrease of the total speech duration with age, mainly in males was seen (Figure 9.5), which is in line with the significant decrease of number of syllables also observed for males. For both genders, speech duration variability decreases with age, but is only significant in males.

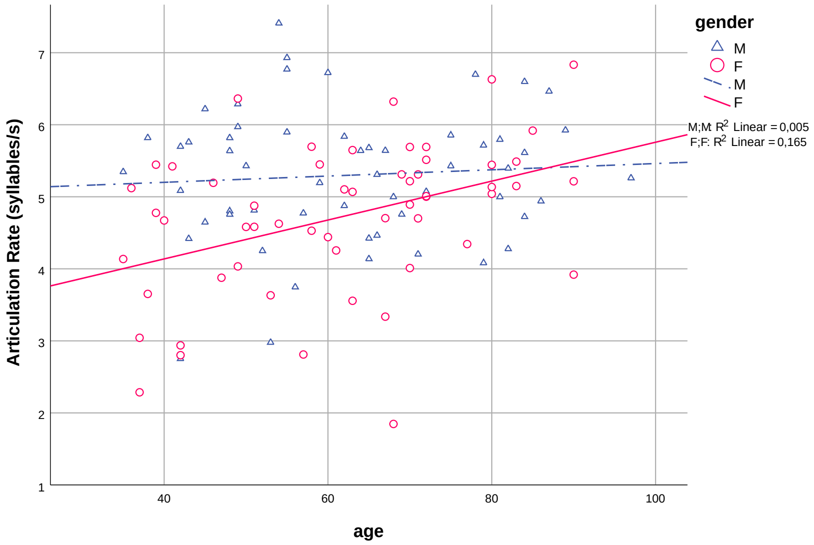


Figure 9.3: Scatterplot and regression lines for articulation rate by age and gender. Each symbol corresponds to one speaker. Solid line and circles: females; dashed line and triangles: males. (adapted from Albuquerque, Valente, Teixeira, Oliveira, and Figueiredo (2021))

9.2.2 Speaking F0

In Figure 9.6 a large age affect in speaking F0 was not observed. Despite non-significant, speaking F0 decreased for females of about 20 Hz between the ages 35 and 100, and increased around 5 Hz for males between the same ages. Thus, in men speaking F0 tended to increase with aging. Conversely, in women speaking F0 tended to slightly decrease with age. As expected, male speakers had a significantly lower speaking F0 when compared to female speakers regardless of age.

9.2.3 Harmonics-to-Noise Ratio

As for HNR, as can be seen in Fig. 9.7, males presented lower values than females, and the regression results confirmed the gender differences. In the case of females, their values tended to be lower with age, while no change in males was observed. The regression line showed a decrease of about 4 dB between the ages 35 and 100 in females, which was significant.

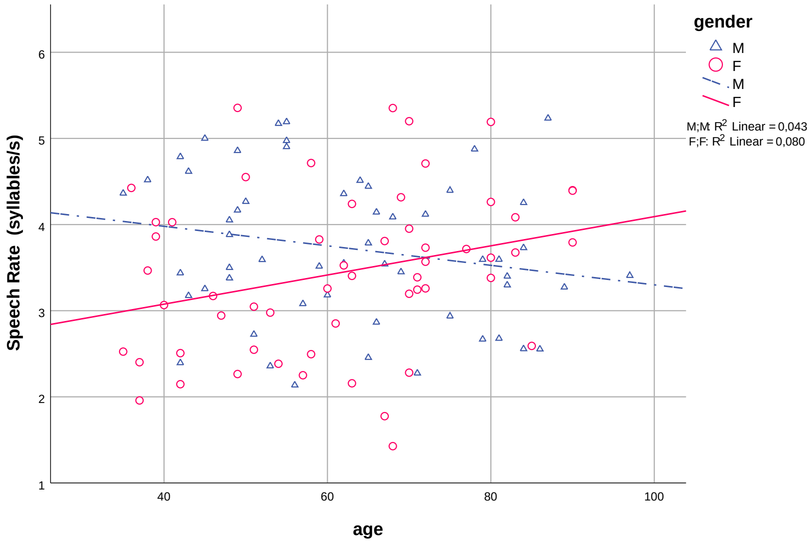


Figure 9.4: Scatterplot and regression lines for speech rate by age and gender. Each symbol corresponds to one speaker. Solid line and circles: females; dashed line and triangles: males. (adapted from Albuquerque, Valente, Teixeira, Oliveira, and Figueiredo (2021))

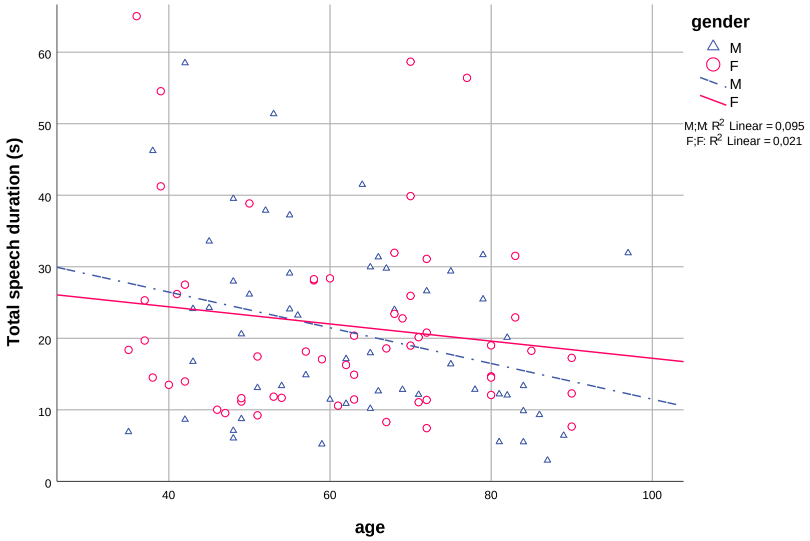


Figure 9.5: Scatterplot and regression lines for total speech duration by age and gender. Each symbol corresponds to one speaker. Solid line and circles: females; dashed line and triangles: males. (adapted from Albuquerque, Valente, Teixeira, Oliveira, and Figueiredo (2021))

9.3 Discussion

This study provides a base to complement previous studies in age effects, mostly at the suprasegmental level. The age effects in rhythmic parameters, speaking F0 and HNR of EP semi-spontaneous speech were explored through a picture description task. Firstly, the older males produced shorter descriptions than younger adults, which may be related to the task nature or indicate differences in linguistic domain (Mortensen, Meyer, & Humphreys, 2006).

The present analyses of rhythmic variation with age are in line with previous findings for other

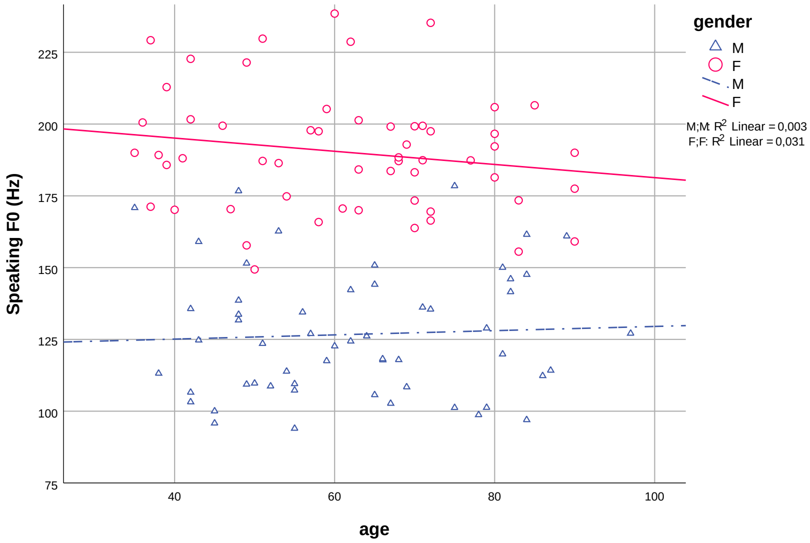


Figure 9.6: Scatterplot and regression lines for Speaking F0 by age and gender. Each symbol corresponds to one speaker. Solid line and circles: females; dashed line and triangles: males. (adapted from Albuquerque, Valente, Teixeira, Oliveira, and Figueiredo (2021))

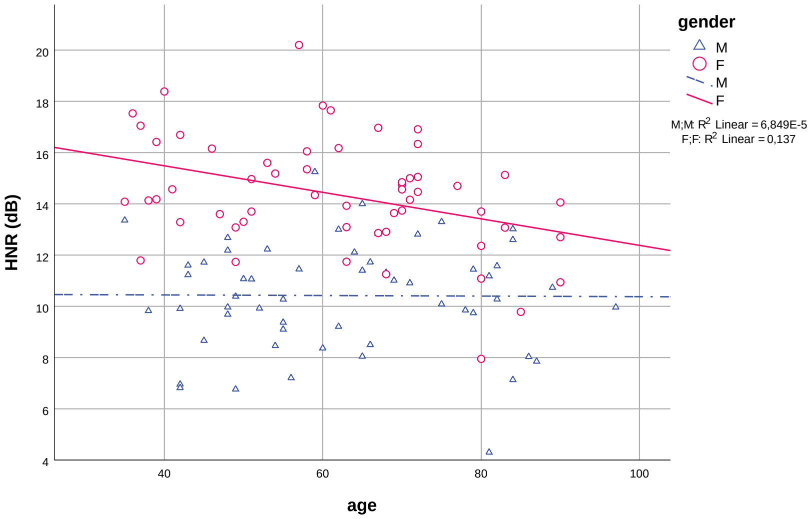


Figure 9.7: Scatterplot and regression lines for HNR by age and gender. Each symbol corresponds to one speaker. Solid line and circles: females; dashed line and triangles: males. (adapted from Albuquerque, Valente, Teixeira, Oliveira, and Figueiredo (2021))

languages, mostly for men (Hartman & Danhauer, 1976; Steffens, 2011), where older men presented significantly more pause time and longer pause duration when compared with the young male speakers. Although men spoke faster than women, as in Jacewicz et al. (2010) and Verhoeven, De Pauw, and Kloots (2004), this difference decreased with age. The faster articulation rate in older women is not in agreement with the general trend (Hazan et al., 2018; Linville, 2001), although some studies (Brückl & Sendlmeier, 2003; Jacewicz et al., 2009, 2010) refer no age-related differences. Furthermore, a longitudinal study of spontaneous speech (Gerstenberg et al., 2018) also reported that articulation rate tends to increase in French speakers with age, presumably due to language specific effects or due to age-related changes in the cardiovascular system (i.e., older speakers may inhale more often, but may compensate for the reduced respiratory capacities by an increase in articulation rate to maintain information density) (Gerstenberg et al., 2018). Also, Brückl and Sendlmeier (2003) suggest that the duration of speech pauses is a better indicator of age than the articulation rate.

Additionally, for males, the results indicated that the total speech duration decreases with age, despite the fact that the mean speech duration of each speech interval only tends to decrease slightly. However, the speech duration variability significantly decreased with age, which may indicate that older males tend to perform speech intervals with a similar duration.

Regarding speaking F0, the data tend to confirm the claim in the literature that F0 slightly decreases in females with age (Goy et al., 2013; Guimarães & Abberton, 2005; Hazan et al., 2018; Morgan & Rastatter, 1986; Titze, 1994; Winkler, 2004), which has been attributed to the endocrinological changes that occur after menopause (Linville, 2001; Schötz, 2006; Sebastian et al., 2012). For males the speaking F0 appears to remain stable with age, based on the regression lines. Nonetheless, the analysis of the same data by age groups (see Figure 9.8) revealed that speaking F0 decreases until the age group [65-79] and starts to rise after that age, with an increase of about 10 Hz in the older group, which presented the highest F0 mean. Thus, a non-linear variation of speaking F0 with age, which has been reported in other studies (Titze, 1994), was also observed in the data. The increase of speaking F0 in older males may be associated with the muscle atrophy, or with an increase in stiffness of vocal folds tissue with aging (Higgins & Saxman, 1991; Linville, 2001; Sebastian et al., 2012). Thus, speaking F0 tends to converge across genders as age increases.

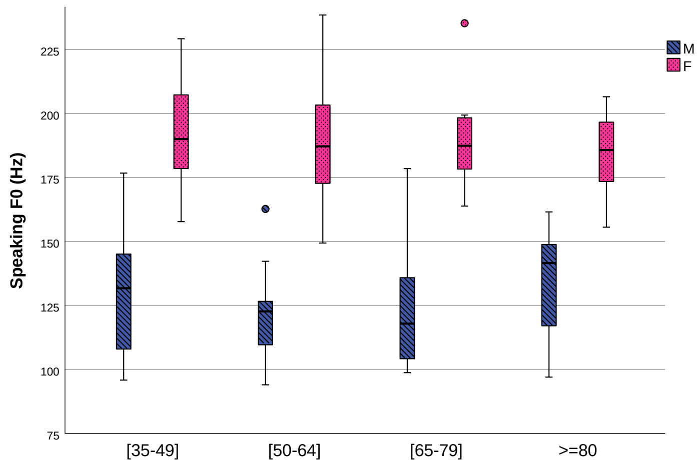


Figure 9.8: The age effect on speaking F0 for both genders (by age groups).

Lastly, females exhibit higher values of HNR than males, as in most studies (Ambreen, Bashir, Tarar, & Kausar, 2019; Dehqan et al., 2013; Goy et al., 2013; Schötz, 2006). The significant decrease of HNR with age in females is also consistent with past studies (Dehqan et al., 2013; Ferrand, 2002; Xue & Deliyski, 2001), and reflects more additive noise in the voiced signal (Dehqan et al., 2013; Ferrand, 2002; Schötz, 2006). According to Mueller (1997), a lower HNR would be predicted because hoarseness tends to be more prevalent in the voices of older speakers due to decrements in laryngeal function, such as ossification of cartilage, degeneration of muscle, connective tissue, and neural tissue, as well as respiratory changes (Dehqan et al., 2013; Mueller, 1997; Titze, 1994). For males, almost no age-related variation was found in HNR, as in Ambreen et al. (2019), Dunashova (2021) and Goy et al. (2013). However, these results should be considered with caution, since the type of speech sample used in the present research did not consist of sustained vowels. Still, a longitudinal study (Dunashova, 2021), using reading samples of one male speaker at different ages (age 59 and age 74), found no changes in HNR with age.

Conclusions on the effects of aging on spontaneous speech should be drawn with caution due to differences in the recording environment, and because of the automatic extraction procedures. Although not all labeled syllables were manually verified, they were obtained using the same procedure for all speakers.

ASSOCIATION BETWEEN ACOUSTIC SPEECH FEATURES AND NON-SEVERE LEVELS OF MOOD SYMPTOMS ACROSS LIFESPAN

The main body of this chapter is based on the study published in the journal PLOS ONE. [1].

Several studies investigated the acoustic effects of diagnosed anxiety and depression. Anxiety and depression are not characteristics of the typical aging process, but minimal or mild symptoms can appear and evolve with age. However, the knowledge about the association between speech and anxiety or depression is scarce for minimal/mild symptoms, typical of healthy aging. As longevity and aging are still a new phenomenon worldwide, posing also several clinical challenges, it is important to improve the understanding of the impact of non-severe mood symptoms on acoustic features across lifespan.

Therefore, the acknowledgment that different acoustic features could be associated with depressive and/or anxiety symptoms lead to the exploration of this relationship in a sample composed by adult participants of different ages. The aim of this study is to determine if variations in segmental and suprasegmental acoustic features have corresponding alterations in anxiety or depression symptoms in adult population considering age. Thus, the present study intends to 1) analyze the association between the acoustic parameters of vowels with depressive and anxiety symptoms; 2) analyze the association between suprasegmental characteristics of semi-spontaneous speech (e.g., rhythmic measures, speaking F0 and HNR) with anxiety and depressive symptoms.

10.1 Method

10.1.1 Participants and tasks

The EP speech database used in this study have previously been described in Chapters 4 and 5. Two different speech tasks (reading vowels in disyllabic words and describing the “Cookie Theft picture”) were performed by 112 individuals aged 35-97.

To assess anxiety and depression symptoms, the Portuguese version of the self-report questionnaire HADS was used (as described in Section 4.4 (Instruments)). Following the cut-offs, a score of 0 – 7 for each subscale could be regarded as being without anxiety or depression symptoms (i.e., 7 is the maximum value for the normal range) (Pais-Ribeiro et al., 2007).

The corpus consisted of two types of parameters: the first refers to segmental and the second to suprasegmental acoustic measures. The speech corpus for segmental analysis consisted of 28 disyllabic

[1] Albuquerque, L., Valente, A. R. S., Teixeira, A., Figueiredo, D., Sa-Couto, P., & Oliveira, C. (2021). Association between acoustic speech features and non-severe levels of anxiety and depression symptoms across lifespan. *PLOS ONE*, 16(4), e0248842.

words, with the EP vowels [i], [e], [ɛ], [a], [o], [ɔ] and [u] in stressed position, mostly composed by a **CV**.CV sequence. The speech corpus for suprasegmental analysis consisted of the "Cookie Theft picture" (Goodglass & Kaplan, 1983) description. For more details about the corpus see Section 4.5.

A set of 19 parameters were extracted from the recording data. Segmental parameters were defined in Table 10.1. Suprasegmental parameters (i.e., rhythmic measures, speaking F0 and HNR) were defined in Table 9.1 (Chapter 9). The segmentation of the data set and the extraction procedures were presented for segmental and suprasegmental data in Chapter 5.

Table 10.1: Description of the segmental parameters used. (adapted from Albuquerque, Valente, Teixeira, Figueiredo, et al. (2021)

Segmental parameters	**Description**
Vowel Fundamental Frequency (Hz)	Median of vibrations per second of the vocal folds on stressed vowels (F0)
Vowel Formant Frequencies (Hz)	The resonance frequencies of the vocal tract (F1 and F2) on stressed vowels
Vowel duration (s)	Mean duration of all EP stressed oral vowels

10.1.2 Statistical analysis

The statistical analysis was conducted with the SPSS software package (SPSS 25.0). The segmental measures (F0, F1, F2 and duration) were obtained for each vowel and, afterwards, median of repetitions was obtained for each vowel type and speaker. F0, F1, F2 and duration mean for stressed vowels were also calculated. The suprasegmental measures (presented in Table 9.1) were also incorporated.

Descriptive data for HADS-A and HADS-D were obtained through the calculation of mean and standard deviation by age (both in a categorical and continuous format), and gender. A two-way ANOVA was performed, including the interaction term between age group and gender. The variance homogeneity (Levene test) and the normality of residuals (by using inspection of QQ plot) were verified. Additionally, descriptive data for segmental and suprasegmental acoustic parameters were reported in mean and standard deviation by gender and HADS-A or HADS-D mood symptoms classification (≤ 7 versus > 7, respectively). Adopting the intensity of change used by Özseven et al. (2018), in the comparison of neutral reading and anxious reading/spontaneous speech, which considers that a high increase is superior to 10% and a high decrease exceeds -10%, the differences between speakers without anxiety/depression symptoms and speakers with mood symptomatology were analyzed by gender.

To explore and model the relationship between all acoustic variables and the scores of mood symptoms (either HADS-A or HADS-D), a multiple linear regression model was developed with non-highly correlated acoustic variables as independent variables (defined as multivariable model). Then the regression models were adjusted for age (continuous) and gender (defined as adjusted model). The assumptions of residuals normality (QQ plot inspection) and homoscedasticity (scatterplot of residuals versus predicted values) were verified. The multicollinearity between independent variables were evaluated by Pearson correlation. Correlation values superior to 0.70 (in module) were considered highly correlated. Acoustic variables that presented a very large (> 0.7) magnitude of correlation (Hopkins, 2002), meaning that they measure the same behaviour and present a similar contribution to the model (Gillam, Logan, & Pearson, 2009), were excluded from the analysis. So, only the acoustic variables vowels F0, vowel duration, vowels F2, total speech duration, total pause duration, speech rate, percent pause time and HNR are included. Due to multiple testing, resulting from the regression models (four models at total), the significant level used was 0.0125.

10.2 Results

First, this section presents the sample characterization in terms of HADS-A and HADS-D scores by gender and age group. Secondly, the association of HADS-A and HADS-D with acoustic parameters are presented.

10.2.1 Sample characterization concerning mood measures

Table 10.2 presents the sample characterization concerning demographic variables and mood measures. Concerning age and gender, the sample is almost balanced. Regarding mood measures, for HADS-A and HADS-D, the number of participants with and without presence of anxiety or depression symptoms is unbalanced (26.8% and 10.7%, respectively) and non-severe (HADS-A: 5.4 ± 2.9 and HADS-D: 4.2 ± 2.7, respectively). HADS-A and HADS-D mean score by age group and gender are also presented in Table 10.2. Figs 10.1 and 10.2 represent the age effect on HADS-A and HADS-D, respectively.

Table 10.2: Sample characterization concerning demographic variables and mood measures. (adapted from Albuquerque, Valente, Teixeira, Figueiredo, et al. (2021))

Variables	**ALL (n=112)**	**Female**	**Male**
Age (years; M ± SD)	62.1 ± 15.6	61.6 ± 15.8	62.6 ± 15.5
Gender (n, %)	112 (100%)	56 (50%)	56 (50%)
HADS_A >7 (n, %)	30 (26.8%)	20 (35.7%)	10 (17.9)
HADS_A [0-21] (M ± SD)	5.4 ± 2.9	5.9 ± 3.3	4.8 ± 2.3
by age (years): [35-49]	5.8 ± 2.8	6.5 ± 3.2	5.1 ± 2.4
[50-64]	6.0 ± 3.2	6.9 ± 3.6	5.1 ± 2.6
[65-79]	4.8 ± 2.4	5.3 ± 2.9	4.3 ± 1.7
≥ 80	4.6 ± 2.8	4.4 ± 3.2	4.8 ± 2.5
HADS_D >7 (n, %)	12 (10.7%)	8 (14.3%)	4 (7.1%)
HADS_D [0-21] (M ± SD)	4.2 ± 2.7	4.3 ± 2.8	4.1 ± 2.6
by age (years): [35-49]	3.2 ± 2.1	3.8 ± 2.2	2.7 ± 2.0
[50-64]	3.8 ± 2.5	3.7 ± 2.9	4.0 ± 2.2
[65-79]	4.2 ± 2.6	4.2 ± 2.9	4.3 ± 2.4
≥ 80	6.2 ± 2.9	6.4 ± 3.0	6.1 ± 3.0

Concerning HADS-A (Fig 10.1), there was a tendency for the median values to decrease after the middle age in female participants; in male participants, the age group [65-79] presented the lower median value of HADS-A. Only females of the age groups [35-49] and [50-64] presented part of the boxplot whiskers above the cut-off of 10 (moderate symptoms of anxiety). Although, two-way ANOVA showed no statistical effect of age ($F(3, 104)$=1.618; p=0.190) or gender ($F(1, 104)$=3.039; p=0.084) on HADS-A scores. Additionally, significant interaction between age group and gender for HADS-A ($F(3; 104)$=0.692; p=0.559) was not detected.

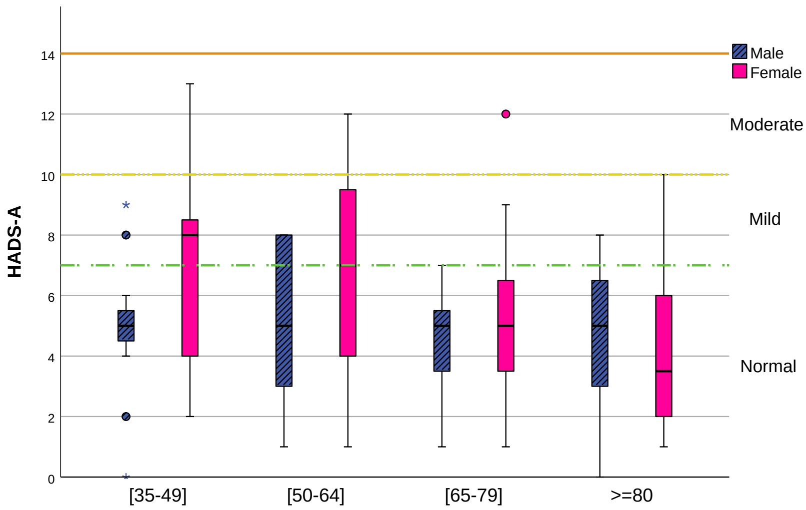


Figure 10.1: Age effect on HADS-A for both genders. (adapted from Albuquerque, Valente, Teixeira, Figueiredo, et al. (2021))

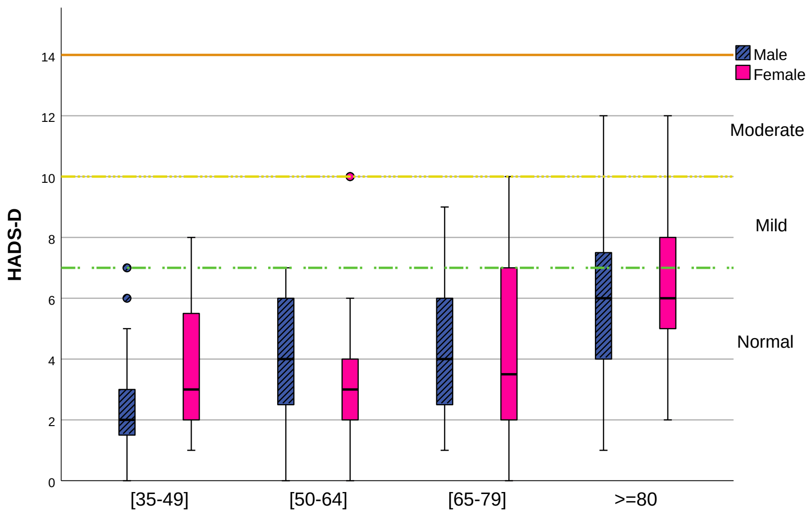


Figure 10.2: Age effect on HADS-D for both genders. (adapted from Albuquerque, Valente, Teixeira, Figueiredo, et al. (2021))

HADS-D (Fig 10.2) tended to increase with age. On male participants HADS-D increased continuously with age and in female participants there was a sharper increase in older age groups. In the older age group, for both genders, the boxplot whiskers achieved the moderate symptoms of depression, but all median values are observed in the normal range. The ANOVA results showed a significant effect of age on HADS-D ($F(3;104)$=6.065; p=0.001), with significant differences between the age group $\geq$ 80 and all the younger groups, but no significant statistical effect of gender ($F(1;104)$=0.275; p=0.601). Additionally, the interaction of age with gender was non-significant ($F(3;104)$=0.470; p=0.704).

10.2.2 Association of HADS-A and HADS-D with acoustic parameters

Considering the division of HADS scores in absence ([0-7]) and presence of symptoms (> 7), the mean and SD values of all acoustic parameters by gender and mood symptoms are presented in Table 10.3.

Table 10.3: Acoustic parameters characterization by HADS sub-scores and gender. (adapted from Albuquerque, Valente, Teixeira, Figueiredo, et al. (2021))

	Variables	Gender	HADS-A		HADS-D	
			≤ 7	**>7**	**≤ 7**	**>7**
			n= 36 (F) + 46 (M)	n= 20 (F) + 10 (M)	n= 48 (F) + 52 (M)	n= 8 (F) + 4 (M)
Segmental	**Vowels F0 (Hz)**	F	192.0 ± 25.8	190.5 ± 20.0	192.6 ± 24.3	184.9 ± 20.3
		M	137.4 ± 30.0	153.1 ± 25.8	137.9 ± 26.6	170.3 ± 53.0
	Vowels F1 (Hz)	F	499.9 ± 34.9	499.5 ± 24.6	500.8 ± 32.8	493.6 ± 21.4
		M	444.9 ± 26.7	451.0 ± 25.9	446.3 ± 27.4	442.1 ± 6.5
	Vowels F2 (Hz)	F	1682.0 ± 62.8	1653.2 ± 56.8	1671.9 ± 63.2	1670.6 ± 56.2
		M	1426.9 ± 59.5	1418.8 ± 62.9	1423.7 ± 57.7	1448.1 ± 88.5
	Vowel duration (s)	F	137.2 ± 28.4	124.8 ± 19.5	130.5 ± 23.1	146.2 ± 39.1
		M	127.9 ± 23.1	126.6 ± 24.6	126.5 ± 23.0	142.8 ± 21.9
Suprasegmental	**Total speech duration (s)**	F	23.6 ± 14.5	18.5 ± 9.9	22.7 ± 13.9	16.6 ± 5.8
		M	20.8 ± 11.8	20.8 ± 16.7	21.5 ± 12.8	11.5 ± 4.5
	Total pause duration (s)	F	7.6 ± 5.4	9.1 ± 7.0	7.7 ± 5.3	10.7 ± 9.2
		M	8.6 ± 5.2	7.8 ± 5.3	8.5 ± 5.2	6.8 ± 5.3
	Total recording duration (s)	F	31.2 ± 17.5	27.6 ± 15.2	30.4 ± 17.1	27.3 ± 14.2
		M	29.4 ± 15.3	28.6 ± 20.6	30.1 ± 16.3	18.3 ± 9.6
	Percent pause time (%)	F	24.4 ± 12.6	31.1 ± 11.3	25.9 ± 11.2	32.0 ± 18.5
		M	29.6 ± 10.9	28.9 ± 12.7	29.2 ± 11.2	33.0 ± 10.8
	Speech pause ratio	F	8.9 ± 25.1	2.7 ± 1.5	6.8 ± 21.6	5.9 ± 9.9
		M	2.8 ± 1.4	3.0 ± 1.5	2.9 ± 1.4	2.3 ± 1.1
	Number of pauses	F	8.2 ± 5.8	8.0 ± 4.7	8.2 ± 5.5	8.1 ± 4.9
		M	8.9 ± 5.0	8.9 ± 7.6	9.2 ± 5.5	5.3 ± 3.7
	Mean pause duration (s)	F	1.0 ± 0.6	1.2 ± 0.5	1.1 ± 0.6	1.1 ± 0.7
		M	1.0 ± 0.4	1.0 ± 0.5	1.0 ± 0.4	1.3 ± 0.2
	Mean speech duration (s)	F	3.1 ± 3.0	2.2 ± 0.7	2.9 ± 2.6	2.1 ± 0.8
		M	2.1 ± 0.7	2.1 ± 0.8	2.1 ± 0.7	2.1 ± 0.5
	Pause variability (s)	F	0.6 ± 0.6	0.7 ± 0.6	0.6 ± 0.5	0.9 ± 0.9
		M	0.6 ± 0.5	0.6 ± 0.7	0.6 ± 0.6	0.6 ± 0.5
	Speech variability (s)	F	1.5 ± 0.7	1.4 ± 0.6	1.5 ± 0.7	1.1 ± 0.5
		M	1.2 ± 0.6	1.3 ± 0.4	1.2 ± 0.5	1.1 ± 0.3
	Number of syllables	F	111.0 ± 76.8	89.4 ± 50.1	107.8 ± 72.4	76.0 ± 31.5
		M	108.9 ± 61.1	98.0 ± 68.5	110.5 ± 62.6	60.3 ± 26.0
	Speech rate (syllables/s)	F	3.5 ± 1.0	3.3 ± 0.8	3.5 ± 0.9	3.1 ± 1.0
		M	3.8 ± 0.9	3.6 ± 0.7	3.7 ± 0.9	3.4 ± 0.3
	Articulation rate (syllables/s)	F	4.7 ± 1.2	4.8 ± 0.8	4.8 ± 1.1	4.5 ± 0.9
		M	5.3 ± 0.9	5.1 ± 1.1	5.3 ± 1.0	5.1 ± 0.3
	Speaking F0 (Hz)	F	190.5 ± 21.0	189.6 ± 20.6	191.6 ± 20.8	181.8 ± 19.1
		M	123.5 ± 19.6	141.9 ± 26.5	124.9 ± 20.7	151.5 ± 24.9
	HNR (dB)	F	14.4 ± 2.2	14.3 ± 2.3	14.4 ± 2.1	14.0 ± 2.9
		M	10.3 ± 2.1	10.8 ± 2.3	10.3 ± 2.1	12.2 ± 1.0

F: female; M: male

Table 10.3 was analyzed based on the intensity and direction of change of each acoustic parameter, and changes higher than 10% (Özseven et al., 2018) are reported next. Considering HADS-A, in the group of female speakers with anxiety symptoms, an increase occurred in total pause duration, percent pause time, mean pause duration and pause variability; a decrease arose in total speech duration, total recording duration, speech pause ratio, mean speech duration and number of syllables. In male speakers, speaking F0 and vowels F0 were the acoustic variables that presented an increase higher than 10%, while a decrease superior to 10% only occurred in the number of syllables.

Regarding depressive symptoms, for females, an increase was observed in vowel duration, total pause duration, percent pause time and pause variability. A decrease occurred in total speech duration, total recording duration, speech pause ratio, mean speech duration, speech variability, number of syllables and speech rate. For males, the acoustic variables that presented an increase higher than 10% were vowels F0, vowels duration, percent pause time, mean pause duration, speaking F0 and HNR. A decrease superior to 10% occurred in the acoustic variables total speech duration, total pause duration, total recording duration, speech pause ratio, number of pauses and number of syllables.

For a more in-depth analysis of the association between the independent variables (i.e., acoustic variables: vowels F0, vowel duration, vowels F2, total speech duration, total pause duration, speech rate, percent pause time and HNR) and the dependent variable (HADS-A or HADS-D scores), a multiple linear regression model was applied and adjusted by the influence of age and gender. In Table 10.4 the multiple regression model results for HADS-A and HADS-D are presented. Although no significant gender differences were observed for both HADS subscales (see Fig 10.1 and Fig 10.2), and only HADS-D presented significant age differences, this approach is justified by the influence of these demographic variables on anxiety/depressive symptoms in other studies in this field (Bromet et al., 2005; Girgus, Yang, & Ferri, 2017; Jorm, 2000; Kessler et al., 2010; Kuehner, 2017; Salk, Hyde, & Abramson, 2017; Van de Velde, Bracke, & Levecque, 2010). For HADS-A, none of the acoustic variables considered presented a significant effect in both models.

Table 10.4: Results of multiple regression model for HADS-A and HADS-D. (adapted from Albuquerque, Valente, Teixeira, Figueiredo, et al. (2021))

Variables	HADS-A				HADS-D			
	Multivariable Model		Adjusted Model		Multivariable Model		Adjusted Model	
	Coeff. 98.75% CI	*p-value*	Coeff. 98.75% CI	*p-value*	Coeff. 98.75% CI	*p-value*	Coeff. 98.75% CI	*p-value*
Constant	4.765 [-7.504 – 17.034]	*0.326*	18.903 [-0.597 – 38.404]	*0.015*	5.108 [-5.834 – 16.049]	*0.238*	6.896 [-10.483 – 24.276]	*0.315*
Vowels F0 (Hz)	-0.002 [-0.029 – 0.025]	*0.869*	-0.008 [-0.035 – 0.019]	*0.438*	0.002 [-0.022 – 0.025]	*0.871*	0.002 [-0.022 – 0.026]	*0.853*
Vowel duration (s)	-0.021 [-0.049 – 0.008]	*0.067*	-0.015 [-0.046 – 0.017]	*0.237*	0.026 [0.00 – 0.051]	*0.011 **	0.012 [-0.016 – 0.040]	*0.270*
Vowels F2 (Hz)	0.000 [-0.007 – 0.008]	*0.927*	-0.008 [-0.019 – 0.004]	*0.092*	-0.001 [-0.008 – 0.005]	*0.611*	-0.003 [-0.013 – 0.007]	*0.455*
Total speech duration (s)	-0.023 [-0.131 – 0.086]	*0.596*	-0.023 [-0.131 – 0.084]	*0.582*	-0.130 [-0.227 – -0.033]	*0.001 **	-0.112 [-0.208 – -0.016]	*0.004 **
Total pause duration (s)	-0.005 [-0.302 – 0.292]	*0.969*	-0.010 [0.304 – 0.283]	*0.929*	0.284 [0.019 – 0.549]	*0.008 **	0.249 [-0.012 – 0.510]	*0.017*
Speech rate (syllables/s)	-0.042 [-1.018 – 0.934]	*0.913*	-0.126 [-1.126 – 0.873]	*0.748*	-0.059 [-0.929 – 0.811]	*0.863*	-0.286 [-1.177 – 0.605]	*0.417*
Percent pause time (%)	0.042 [-0.094 – 0.178]	*0.437*	0.053 [-0.080 – 0.187]	*0.314*	-0.086 [-0.208 – 0.035]	*0.074*	-0.090 [-0.209 – 0.029]	*0.057*
HNR (dB)	0.214 [-0.116 – 0.544]	*0.103*	0.119 [-0.217 – 0.455]	*0.371*	0.052 [-0.243 – 0.347]	*0.655*	0.074 [-0.226 – 0.373]	*0.533*
Age	–	–	-0.028 [-0.080 – 0.024]	*0.177*	–	–	0.047 [0.001 – 0.094]	*0.010 **
Gender (Female)	–	–	3.093 [-0.297 – 6.483]	*0.022*	–	–	0.439 [-2.582 – 3.460]	*0.712*

Coeff.: Unstandardized Coefficient; * $p<0.0125$

For depression symptoms, expressed by the HADS-D scores, vowel duration, total speech duration and total pause duration presented significant effects. However, in the adjusted model only total speech duration maintained the significant effect, along with age. In the adjusted model, age was also significantly associated with HADS-D.

Fig 10.3 demonstrates the association of the depression symptoms scores and the total speech duration and age. The increase of depressive symptoms is related to the total speech duration decrease and to age increase.

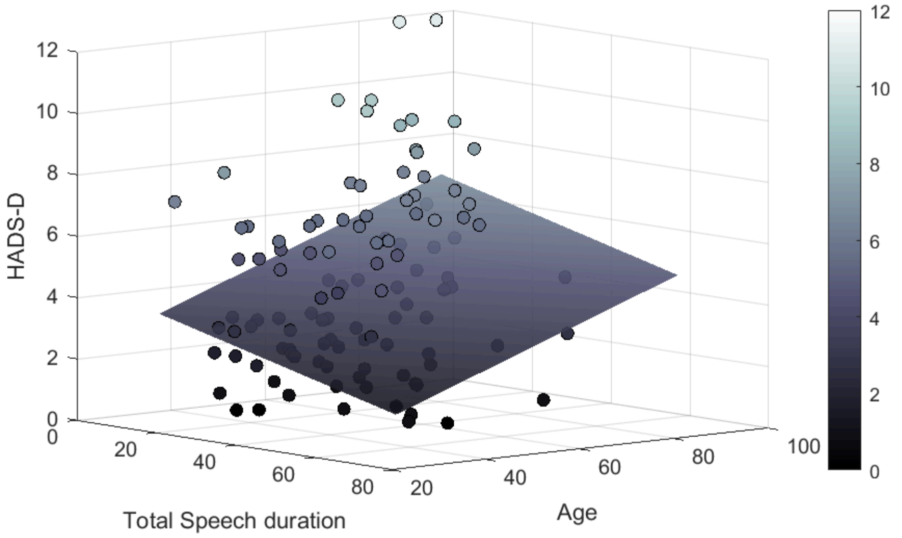


Figure 10.3: Relationship between total speech duration, age and HADS-D. (adapted from Albuquerque, Valente, Teixeira, Figueiredo, et al. (2021))

10.3 Discussion

The present study aimed to analyze the relationship between the scores of the HADS questionnaire and the segmental acoustic parameters (e.g., F0, F1, F2 and duration of stressed vowels) and also the suprasegmental measures obtained in a sample of 112 individuals (aged 35 to 97) with non-severe mood symptoms. The aim of the study was achieved considering the general alignment of the present results with previously reported research related with mood diagnosed disorders.

Regarding anxiety symptoms, there are no acoustic variables that presented a significant association with HADS-A scores. The independent variables used to develop the multiple linear regression do not present a high increase/decrease difference between participants without anxious symptoms and participants with anxious symptoms (see Table 10.3). In fact, the majority of those differences were below 10% (mainly in males), which can be an explanation for the non-significance observed in the multiple linear regression. The authors can argue that this minor difference could not be sufficient to make the acoustic variables sensitive to sub-clinical anxiety symptoms. Additionally, the observed tendency to higher HADS-A values in younger females has been reported in other studies (Girgus et al., 2017; Kuehner, 2017; Salk et al., 2017; Van de Velde et al., 2010), due to interactions between behaviors, internal gender characteristics and stressors (Girgus & Yang, 2015).

For depression symptomatology, this study presents significant results for both segmental and suprasegmental levels. At segmental level, vowel duration presents a significant effect of the depressive symptoms, meaning that vowel duration increases as depressive symptoms increase. The significant effect of depressive symptoms in vowel duration found in the present study, which analyzed a sample mostly constituted by speakers with non-severe depressive symptoms, also highlight the importance of segment duration in the identification of mood signs. The current results are in line with Trevino et al. (2011) that, through the use of phone-duration measures instead of global measures of speech rate, found significant positive correlations between the duration of some vowels with the worsening of depression. To reinforce the results obtained in the present study, the findings of Alghowinem et al. (2012) and Hönig et al. (2014) can also be reported, which concluded that syllable duration (in average) was significantly higher in the group of depressed individuals.

At suprasegmental level, first, the total pause duration increases with more depressive symptoms, and the total speech duration presents an inverse trend. Mundt et al. (2007, 2012) revealed that great depression symptoms result in more and longer pauses, which was also reflected in a higher total pause time, as occurred in the present research. Conversely, in both studies (Mundt et al., 2007, 2012) more time was needed to deliver the message (i.e., more total recording duration due to longer pause time during speech production). However, other studies have reported that speakers with depressive symptoms exhibit a decrease in speech time or in verbal productivity (Esposito, Esposito, Likforman-Sulem, Maldonato, & Vinciarelli, 2016; Hall, Harrigan, & Rosenthal, 1995; Klumpp & Deldin, 2010). That is, these speakers tend to produce fewer words (Klumpp & Deldin, 2010) and tend to decrease the phonation time (i.e., utterances are shorter in duration and are less numerous) (Esposito et al., 2016). Results concerning the increase in total pause duration and the decrease in total speech duration (the one that maintained the significant effect on the adjusted model) could be considered an index of psychomotor retardation or lower cognitive function, and affect the amount of information content to be communicated (Esposito et al., 2016). In the present study, the total speech duration decreases in speakers with more depressive symptoms (a difference of -32.6% for participants with depressive symptoms (considering both genders)), due to the fact that semi-spontaneous speech production is more cognitively demanding in comparison with automatic speech/reading tasks, requiring preparation, word selection and higher motor articulatory control (Esposito et al., 2016; Mundt et al., 2007). The increase in total pause duration in participants with more depressive symptoms could suggest more efforts in communication planning and higher cognitive elaboration time (Esposito et al., 2016). The current results (i.e., significant effect of HADS-D on total pause duration and total speech duration) also highlight the importance of rhythmic measures assessed in spontaneous speech for depression symptoms recognition.

Additionally, although the acoustic variable number of syllables has not entered in the regression model, due to the high correlation with total speech duration, in the descriptive data a decrease of more than 10% in the number of syllables was observed for both genders with score > 7 on HADS-D (see Table 10.3). The number of syllables decrease may be related with the total speech duration decrease and the total pause duration increase with the depression worsening.

Although speech rate in semi-spontaneous speech does not present a significant effect in depressive symptoms, considering that vowels constitute the syllable nucleus, an increase in the time needed to produce a vowel could contribute to a decrease in syllable production per time unit (K. R. Scherer, 1986) and, consequently, a decrease in speech rate in reading task. Speech rate is referred in the literature as one of the most strongly associated acoustic features with depression status (Cummins, Sethu, et al., 2015) and also one of the first symptoms of depressive disorders, observable by interlocutors

(Kraepelin, 1921). The literature indicates that individuals with more depressive symptoms present lower speech rates (Cannizzaro et al., 2004; Ellgring & Scherer, 1996; Mundt et al., 2012; Sobin & Sackeim, 1997; Stassen, Kuny, & Hell, 1998), even in brief sadness induction (Sobin & Alpert, 1999). The sensitivity of speech rate for recovery of depressive symptoms has also been evidenced, as the improvement in symptomatology has a positive influence on speech rate (Alpert et al., 2001; Kuny & Stassen, 1993).

The significant findings mentioned above concerning the rhythmic measures (i.e., total pause duration) and vowel duration do not maintain the effects on the adjusted model by age and gender, which provides evidence of a greater influence of age in mood symptoms. Age is the demographic variable that presents a significant effect on the depressive symptoms assessed by HADS-D. Depression symptoms presented statistically significant higher values in older adults, which is in accordance with studies developed in low-income countries (Bromet et al., 2005; Kessler et al., 2010). On epidemiologic studies in Western countries, the rate of depression decreases with age, which is the opposite performance of depression mean values across age in the present study. Balabanova and McKee (2002), and Bobak et al. (2004) suggest that high numbers of depression symptoms in older ages could mirror high levels of poverty or poor physical health. Bromet et al. (2005) also suggest that an increase of depression in older adults could reflect negative changes in social support and in subjective health. The largest European research study on aging (DO-HEALTH) concludes that older individuals in Portugal present low levels of cognitive and physical health compared to other European countries (Bischoff-Ferrari, 2019; Schietzel et al., 2022). In fact, only 9% of the Portuguese sample are considered healthy, much lower than the 58% from Austria, 51% from Switzerland or 38% from Germany (Schietzel et al., 2022). Schietzel et al. (2022) revealed that different social resources could explain poor health levels, including level of education, values of pensions or ease access to health care.

Several differences found between the present study and previous research could be due to the presence of having participants with absent-to-mild symptoms, whereas most studies also include individuals with severe mood symptoms.

This study presented some limitations. The first limitation is related with the task nature used to extract suprasegmental features, once the older speakers performed smaller descriptions than younger adults. This may be related with task or indicate differences in linguistic domain (Mortensen et al., 2006). Secondly, only a small number of speakers had mild or moderate mood symptoms. Lastly, some results should be considered with caution due to the recording environment and the automatic extraction procedures, considering that labelled syllables were not manually verified, but they were obtained in a standardized way for all speakers. Additionally, while certain acoustic features were found to be associated and important in explaining depression and anxiety symptoms, machine learning models would be needed to determine how important they are predicting these mental health states.

This page intentionally contains only this sentence.

AGE-RELATED CHANGES IN EUROPEAN PORTUGUESE VOWEL PRODUCTION: AN ULTRASOUND STUDY

The main body of this chapter is based on the results presented on the study published in Applied Sciences [1].

For aging speech, there is a limited knowledge regarding the articulatory adjustments underlying the acoustic findings observed in previous studies. In this context, ultrasound imaging is a technology that can be safely used to study static and dynamic features of the articulators (namely the tongue contours), allowing comparisons of physiological differences between older and young adults during speech production.

In order to investigate the age-related articulatory differences in EP vowels, the present study analyzes the tongue configuration of the nine EP oral vowels (isolated context and pseudoword context) produced by 10 female speakers of two different age groups (young and old). Thus, the main purpose of this work is to perform a study of the age-related articulatory differences in EP vowels with US imaging and to investigate normalization procedures. Additionally, since there is a paucity of literature on EP oral vowels production, this study also provides valuable insights for the accurate articulatory description of EP oral vowels with US.

11.1 Method

11.1.1 Speakers and corpus

A subset of 10 healthy EP native female speakers (out of a total of 32 speakers - see Section 6.2) were analyzed: 5 young (between 23 and 32, with a mean age of 27.0) and 5 old (between 59 and 73, with a mean age of 62.4) females. They were selected due to high quality of their US data (which was crucial to the articulatory analysis (Strycharczuk & Scobbie, 2017)), and only one gender was selected to avoid possible confounding effects of gender (Goozée et al., 2005).

Considering the speakers' anatomical characteristics, young (Y) and old (O) female speakers presented similar height (Y=161.80 ± 5.67; O=161.40 ± 3.07), but different weight (Y=58.00 ± 8.46; O=66.60 ± 11.83), which is reflected in the greater occurrence of overweight among older women (Table 11.1). All speakers have 9 years or more of education, with the majority (90%) having completed higher education.

The corpus consisted of all EP oral vowels [i], [e], [ɛ], [a], [o], [ɔ], [u], [ɨ] and [ɐ] in pseudoword context and in isolated context. The stimuli were embedded in a carrier sentence that was repeated 3

[1] Albuquerque, L., Valente, A. R., Barros, F., Teixeira, A., Silva, S., Martins, P., & Oliveira, C. (2022). Exploring the Age Effects on European Portuguese Vowel Production: An Ultrasound Study. *Applied Sciences*, 12(3), 1396.

times. For more details about the corpus see Section 6.4.

Table 11.1: Weight and height of the female speakers. (adapted from Albuquerque, Valente, et al. (2022))

Speaker	Weight (kg)	Height (cm)	BMI*	Weight status
Y1	72	158	28.8	Overweight
Y2	62	160	24.2	Normal weight
Y3	50	158	20.0	Normal weight
Y4	57	173	19.0	Normal weight
Y5	49	160	19.1	Normal weight
O1	48	158	19.2	Normal weight
O2	60	167	21.5	Normal weight
O3	80	162	30.5	Obesity class I
O4	67	160	26.2	Overweight
O5	78	160	30.5	Obesity class I

* BMI = $\frac{weight(kg)}{(height)^2(m)}$

11.1.2 Data processing

After acquisition (see Section 6) and synchronization of the US images with the speech audio data, the data obtained were processed in several steps: 1) segmentation of the acoustic data signal and manual revision; 2) US image processing using a U-net network (Ronneberger, Fischer, & Brox, 2015); 3) automatic extraction of the tongue contours in cartesian coordinates from the vowel midpoint; 4) manual verification of the extracted contours; 5) bite plane extraction for each block of vowels; 6) tongue contour rotation to ensure a common referential across blocks; 7) inspection of the US data with a particular emphasis on the [a] sustained vowels, as a quality check of the previous steps; 8) export of US data contours and 9) data pruning of the tongue contours. Each of these steps will be detailed below.

After the automatic segmentation at word and phoneme level using WebMAUS General for Portuguese language (PT) (Kisler et al., 2017), the accuracy of the target vowel boundaries were manually checked in Praat software (Boersma & Weenink, 2012) and corrected if necessary. A total of 22 recordings were discarded (approximately 1.4% of trials) due to problems with the recordings (e.g., the participants misread the target word or the target vowel). After acoustic segmentation, all frames that corresponded to the target vowels were extracted from the US videos using a Python script. A database was devised, containing both vowel contexts, pseudoword (pVCv) and isolated vowels.

The US images were processed and splines were automatically fitted to the surface of the tongue across the vowel's labelled duration using a U-Net adapted from Ronneberger et al. (2015) (already used in US images by Mozaffari and Lee (2019); Zhu, Styler, and Calloway (2019)). This encoder-decoder model was trained with a dataset of 7765 US images of different speakers that were manually annotated by four trained analyzers. In order to virtually increase the amount of training data, data augmentation was also used to train the neural network, and the final training set consisted of 9131 images. Figures 11.1**b** and 11.1**d** show two examples of the automatic tongue contour obtained by the developed method.

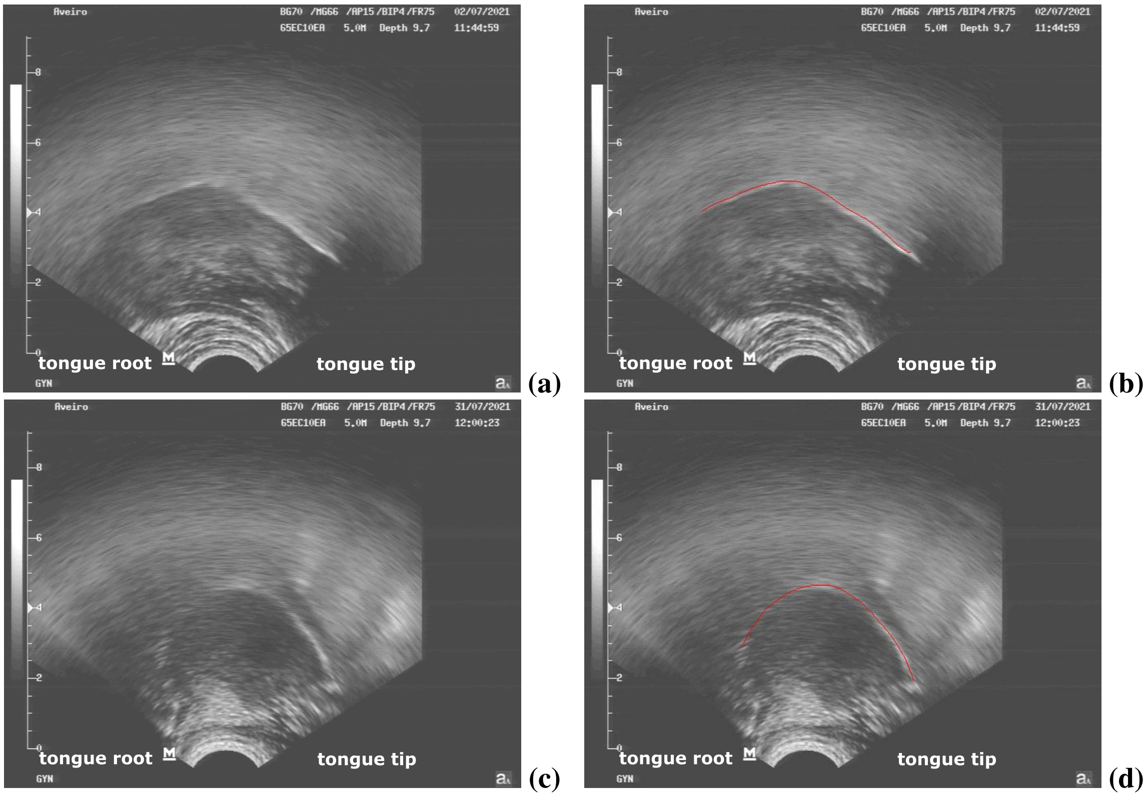


Figure 11.1: Illustrative US images with the automatic tongue spline of the tongue contour for vowel [a] (**a**) (**b**) and vowel [u] (**c**) (**d**). Left: raw US images; Right: US images with automatic tongue spline. (adapted from Albuquerque, Valente, et al. (2022))

For the present study, only the tongue contour corresponding to the temporal midpoint of the vowels was exported, which consistently contained an articulatorily steady part of the vowel (Dokovova et al., 2019). When the vowels presented an even number of frames, both frames were exported. The output data structure contains the cartesian coordinates of the automatic tongue contour of each frame, as a list of points (x,y values in pixel). Given the challenging nature of the images, and to ensure the reliability of the data, the automatic tongue segmentation of the central frames of the vowel occurrences was revised by three annotators with experience in speech production analysis.

2303 splines were obtained through automatic tongue segmentation and 227 images were excluded due to undefined border of the tongue or extremely dark US image. The final number of splines, after manual check, were 2076, which corresponded to 1440 vowels.

The bite plane was extracted for each block of vowels. The bite plane recording which appeared to have the most stable pressing of the tongue was used, and this recording may have been at the beginning or at the end of each randomized block. Each bite plane is represented by the points A and B (see Figure 6.2**b**). For each point A and B the coordinates x and y in pixel were determined to obtain the bite plane line. Nonetheless, for Speaker O5, it was not possible to get one bite plane per block. As the probe orientation was fixed along the session, the tongue contours were rotated based on the available speaker's bite planes. A θ values were, then, calculated, as the angle between the bite plane and the horizontalization of the occlusal plane (see Figure 11.2**a**).

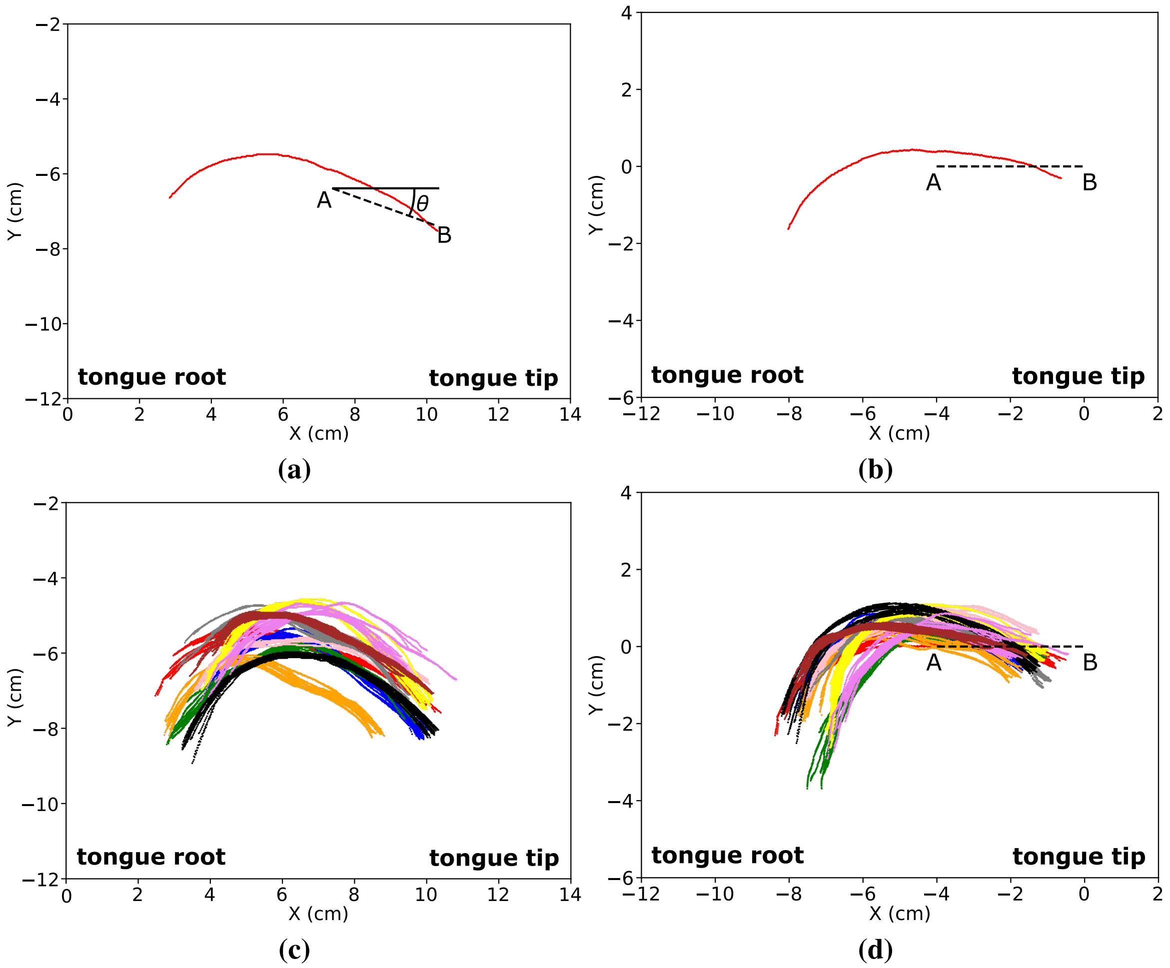


Figure 11.2: Example of unrotated (**a**) and rotated (**b**) tongue tracing of one sustained vowel [a] (top). Tongue contour of each sustained vowel [a] of all speakers unrotated (**a**) and rotated (**a**) (bottom). The dashed line represents the annotation of the bite plate. (adapted from Albuquerque, Valente, et al. (2022))

In order to standardize the rotation of the data across the speakers, for each block of vowels, the tongue contour data were rotated based on the occlusal plane obtained in the corresponding block (Dokovova et al., 2019; Scobbie et al., 2011; Strycharczuk & Scobbie, 2017). As seen in Figure 11.2, the occlusal plane was observed to be parallel to the x axis after rotation (see Figures 11.2**b**). During rotation, pixel to cm conversion was made. Data were exported after rotation and the origin of the coordinates system corresponds to the upper incisors, which are 4 cm from back of the bite plate.

Four frames of each sustained vowel [a] were exported and manually verified. The most complete contour of each [a] was chosen, and the contours were overlapped to inspect the US data. Figure 11.2 presents unrotated (Figure 11.2**c**) and rotated (Figure 11.2**d**) [a] contours of all speakers. Despite the anatomical differences between speakers, the raw coordinate values are partly comparable across the speakers, since the tongue advancement/retraction consists in a distance to the upper incisors. As seen in Figure 11.2, after rotation the tongue contours of the sustained vowels [a] of all speakers tend to be more overlapping.

The tongue contours selected for analysis were revised to detect erroneous/incomplete segmentations mostly due to poor image quality rendering the tongue incomplete. This resulted in the exclusion of 113 contours. Additionally, poor US image quality for some speakers (namely for O2 and O5), vowels (mostly vowel [i]), and/or context (mostly in pVtv context) resulted in less data usable for analysis. Table 11.2 presents the number of tongue contours that are considered for study after the manual check of the US data.

Table 11.2: Number of vowels analyzed per speaker and vowel type after removing some of the data due to poor image quality. (adapted from Albuquerque, Valente, et al. (2022))

	Speaker										
	Y1	**Y2**	**Y3**	**Y4**	**Y5**	**O1**	**O2**	**O3**	**O4**	**O5**	**Total**
[i]	17	18	14	17	17	2	5	6	1	0	*97*
[e]	18	18	17	18	18	17	10	16	8	2	*142*
[ɛ]	18	18	18	18	18	17	16	15	16	7	*161*
[ɨ]	17	18	18	18	17	17	10	17	13	6	*151*
[ɐ]	18	18	18	18	18	16	18	18	16	18	*176*
[a]	18	18	18	17	16	18	14	18	17	17	*171*
[u]	13	18	18	17	18	18	0	14	12	14	*142*
[o]	18	16	14	18	18	18	0	11	12	17	*142*
[ɔ]	18	15	18	18	18	18	0	16	17	17	*155*
Total	*155*	*157*	*153*	*159*	*158*	*141*	*73*	*131*	*112*	*98*	1337

Overall, as expected, the best US images come from sounds where the tongue surfaces are fairly flat and gently curved, such as the central vowels ([a], and [ɐ]). Generally, a less discernible tongue contour occurs at the tip and back regions of the tongue, and vowels that have steep slopes, such as [i] or [u], tend to present worse images. Stone (2005) also explained that the edges perpendicular to the beam will image best and edges more than 50 degrees from perpendicular begin to image poorly. Other studies of vowel production report similar difficulties (Georgeton et al., 2016; Song, 2017).

11.1.3 Articulatory measures and normalization

For each vowel token, two parameters were extracted, namely the y-coordinate of the highest point of the tongue's contour (i.e., tongue height, TH), and the corresponding x-coordinate, that reflects the front back position of the tongue (i.e., tongue advancement, TA). Thus, the TH corresponds to the distance between the bite plane and the highest point of the contour, while TA was measured as the distance to the upper incisors. The TA values are always negative, the higher values (in module) indicate greater tongue retraction and lower values (in module) suggest greater tongue advancement of the highest point of the tongue. When the vowels present two central frames, the mean values of TA and TH were considered.

In order to compare average tongue measures between age groups, it was necessary to adopt an approach towards a normalization of the data based on vocal tract size and US probe orientation. Therefore, the raw measures of vowels are sensitive to anatomical differences between speakers, with expected larger vowel spaces for larger speakers (Strycharczuk & Scobbie, 2017). One potential normalization method would consist of z-score normalization, but in this case a full set of vowels or a normalization in function of the distance to a reference vowel would be required (Strycharczuk & Scobbie, 2017). On the other hand, Zharkova et al. (2011) applied a normalization method based on the assumption that the length of tongue contour correlates with the overall size of the vocal tract, and the proportion of tongue imaged in the ultrasound display is similar across speakers.

Nonetheless, in this study the normalization is independent of the proportion of the tongue imaged for each speaker, but involves the assumption that each distance used as scale factor (TA and TH scale factor) correlates with the overall size of the vocal tract. For TH, the normalization for vocal tract height was carried out on the basis of the distance of the back of the bite plate to the virtual origin of the probe (in mm), for each speaker by block of vowels (see Figure 11.3**a**). The normalization factor for vocal tract length was based on the tongue contour of the sustained vowels [a], because the tongue tip and the larynx are low during the production of this vowel, thus the imaged tongue contour tends to be better (Zharkova et al., 2011). TA normalization factor was based on the distance, in mm, of the upper incisors to the interception between the occlusal plane and the tongue root surface in the production of sustained [a] (mean of all [a] productions), for each speaker (see Figure 11.3**b**).

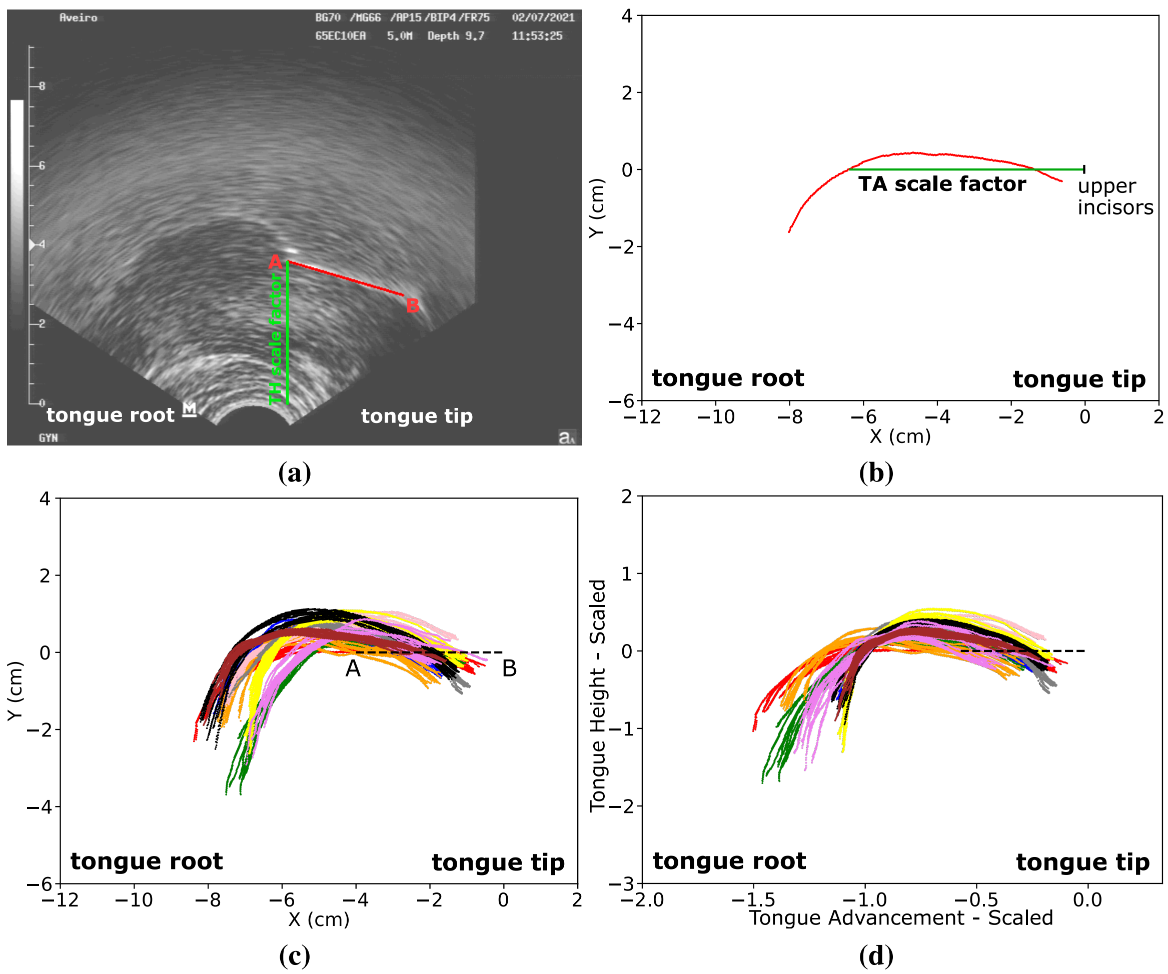


Figure 11.3: Top: Illustration of normalization measures for tongue height (TH) (**a**) and tongue advancement (TA) (**b**). Bottom: Sustained vowels [a] contours of all speakers rotated (**c**) and scaled (**d**). (adapted from Albuquerque, Valente, et al. (2022))

For each speaker, TH and TA normalization was applied, multiplying each raw value of TH and

TA by the corresponding normalization factor. Figures 11.3**c** and 11.3**d** illustrate the normalization procedures.

11.2 Results

In this section a summary of the main findings on EP vowel tongue configurations by age group is presented. Also some considerations about vowel context effect, and inter and intra-speaker differences are reported. Note that, for some older speakers (see Table 11.2) there are no tongue contours available for some vowels, which does not allow a more comprehensive comparison between age groups to be carried out. In this section the scaled US data are used (i.e., TA and TH obtained after normalization). The analysis starts with a brief inspection of the tongue contours. Thus, some examples of tongue contours for all vowels by context are presented in Figure 11.4 for one speaker (speaker Y1).

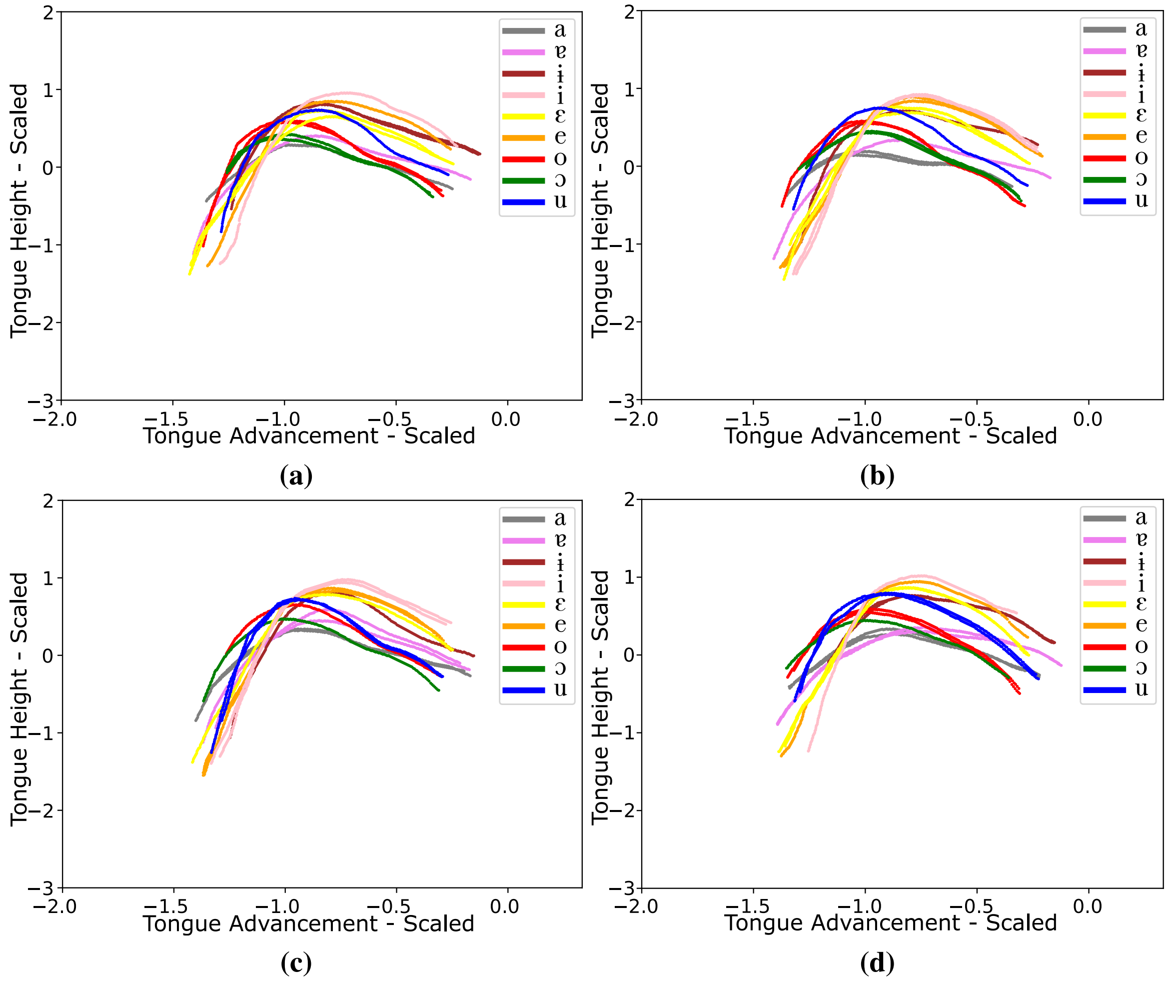
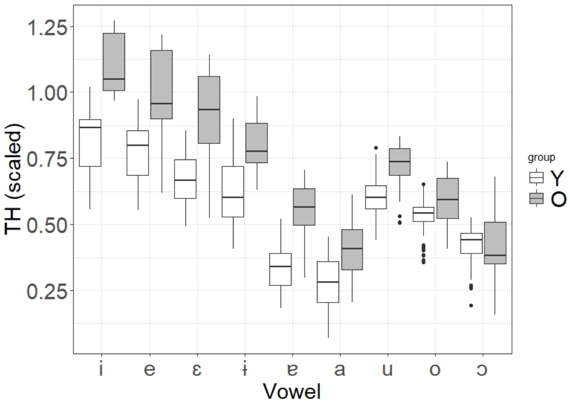
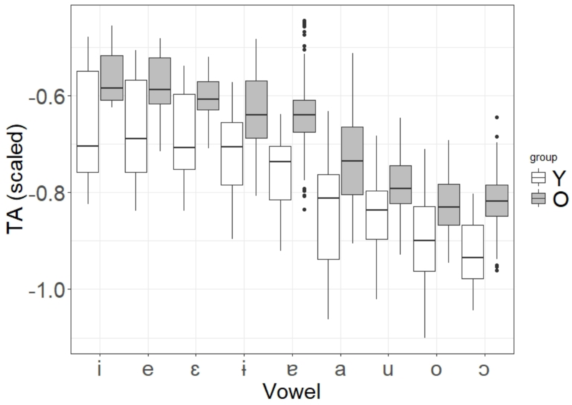


Figure 11.4: Example of one tongue contour by vowel for each context for speaker Y1. (**a**) pVpv context; (**b**) pVtv context; (**c**) pVkv context; (**d**) isolated context. (adapted from Albuquerque, Valente, et al. (2022))

Considering the articulatory measures, Figure 11.5 summarizes the TH (Figure 11.5**a**) and TA (Figure 11.5**b**) values obtained by vowel for each age group. In this first analysis, both contexts (vowels in pVCv sequence and isolated) are considered. In general, concerning TH, the older females showed higher values, except for vowel [ɔ]. In general, the front vowels presented higher TH values, and the vowels [a] and [ɔ] the lower values, mainly for the older females. For young females the central vowels [ɐ] and [a] showed similar TH, while for older females the vowel [ɐ] presented higher TH values than vowel [a]. Also, old females presented identical values of TH for vowel [a] and [ɔ].



(a)



(b)

Figure 11.5: Boxplots of tongue height (TH) (**a**) and tongue advancement (TA) (**b**) values by vowel and age group. (adapted from Albuquerque, Valente, et al. (2022))

Regarding TA, the lowest values (in module) were obtained for the old group. For both groups, the front vowels tend to present the lowest TA, and the vowels [o] and [ɔ] revealed the highest TA values (in module). Variability tends to be higher in TA in comparison with TH values, mainly for young females (i.e., the variability of TA tends to decrease with age).

To complement this initial analysis, the individual variability is also explored through the articulatory vowel space and the vowel articulatory cluster size by speaker and vowel context.

Figure 11.6 represents the articulatory vowel space defined by scaled TA and TH values of the cardinal EP oral vowels ([a], [i] and [u]) for each speaker and vowel context (i.e., pVCv and isolated vowel). Each vowel is represented by the mean TH and TA. As four old females (O1, O2, O4 and O5) presented incomplete data, it was not possible to define the articulatory space of all speakers by vowel context. Furthermore, the young females also showed great variability in the shape of the articulatory vowel space. Concerning the vowel context, it can be observed that articulatory space tends to be smaller when vowels occur in pVCv sequences comparing with isolated vowels, namely for the speakers Y2, Y3, Y5 and O3.

Since the main goal is the study of age effects, in Figure 11.7 the articulatory vowel space of the two age groups is compared, separately for each production context. The articulatory space was defined by the mean of the mean TA and TH values of the cardinal EP oral vowels for each speaker by vowel context.

The vowel articulatory space tends to be different, in shape and size, between the younger and older females in both contexts. Old females produced vowels with higher TH and a notable decrease (in module) of the TA (i.e., the highest point of the tongue contour tended to be more advanced). Consequently these results are in line with some of the observations made above. The vowel articulatory

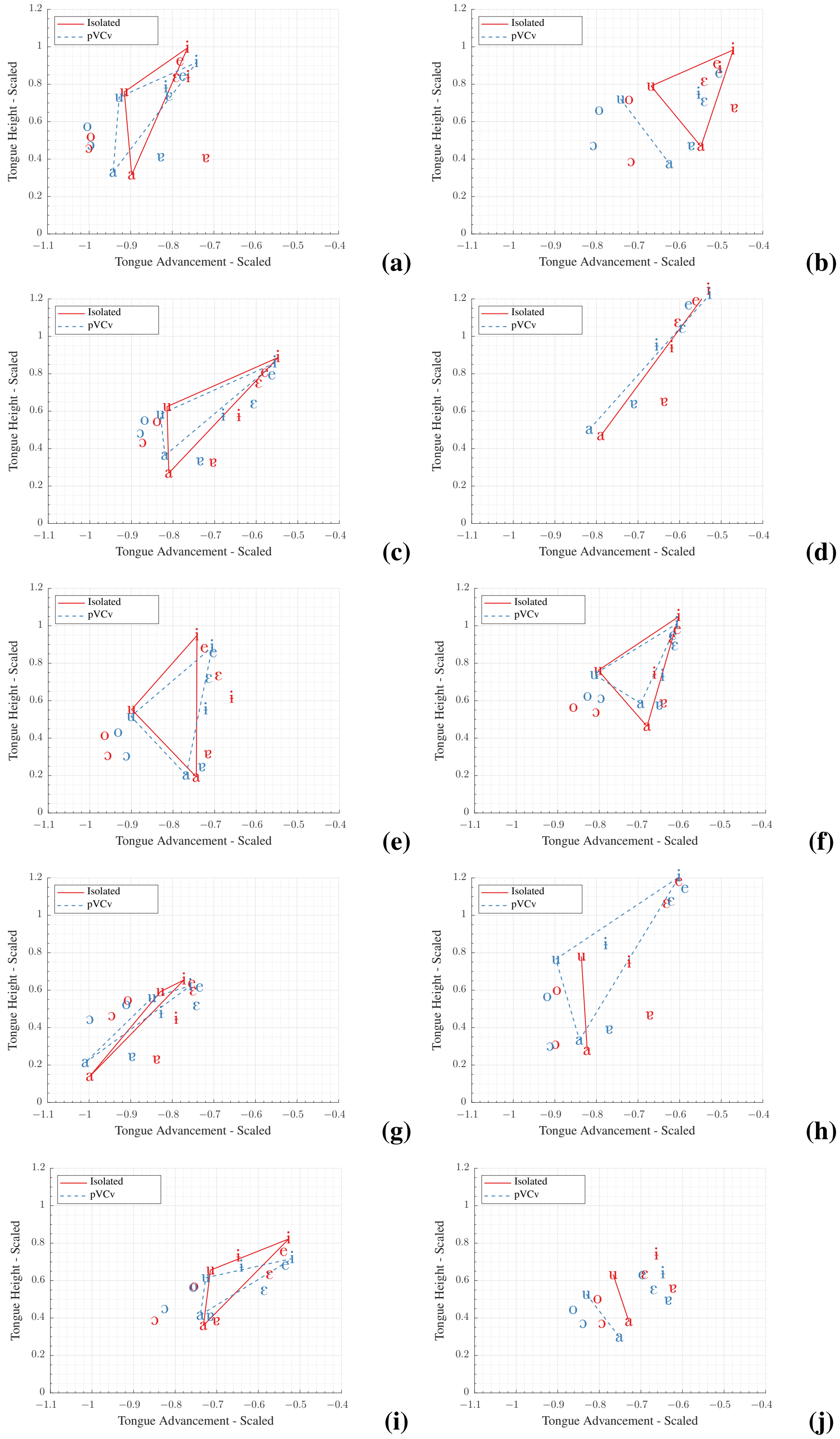


Figure 11.6: Articulatory vowel space of the EP cardinal vowels in isolated (red solid lines) and in pVCv context (blue dashed lines). The remaining vowels are also represented. Left Side: young females (Y1 to Y5) (**a**) (**c**) (**e**) (**g**) (**i**); right side: old females (O1 and O5) (**b**) (**d**) (**f**) (**h**) (**j**). (adapted from Albuquerque, Valente, et al. (2022))

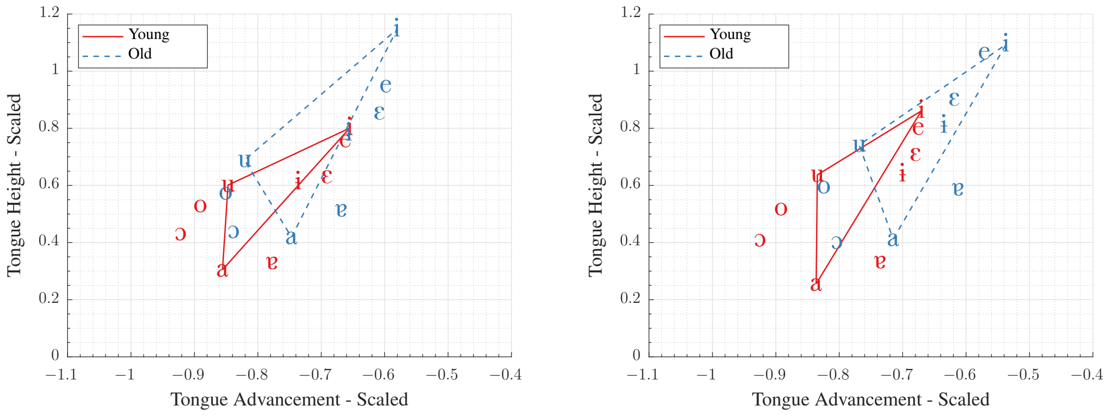


Figure 11.7: Articulatory vowel space of the EP cardinal vowels in pVCv (**a**) and in isolated context (**b**) by age group. The remaining vowels are also represented. (adapted from Albuquerque, Valente, et al. (2022))

space differences observed between both contexts is more pronounced for the young females. That is, the articulatory vowel space area, namely for pVCv sequences, tended to be smaller in the young females. In old females the differences between contexts were smaller, but a tendency to TA decrease (in module) for isolated vowels comparing with vowels produced in pVCv sequences (i.e., isolated vowels tend to be more advanced) was observed.

Plots in Fig. 11.6 only provide the average, but variability of productions is also very important to analyze. Figure 11.8 represents individual productions (i.e., the TH and TA of the total number of occurrences of all EP oral vowels) and information regarding dispersion based on ellipses for two speakers (one young and one old) by vowel context. The ellipses surround values that fall within 2σ of the mean (Escudero et al., 2009). In those graphs a considerable variability between speakers can be observed.

11.3 Discussion

This study contributes to increase knowledge on EP aging speech, providing an articulatory perspective of the effects of age in all oral vowels of the EP based on articulatory measures: TH and TA. The present study extends in many ways the previous pilot research (Albuquerque, Valente, et al., 2022) by reporting articulatory data from more female speakers, and by ensuring an inter-speaker comparison through the application of normalization procedures. Also the automatic method of contours tracing was improved.

In general, the results of this articulatory study reveal that the highest point of the tongue tends to be higher for the older females and more advanced, compared to the younger females (except the TH of the vowel [ɔ]). The increase of the TH is in accordance with the previous acoustic study about aging effects on EP oral vowels (Section 7). In other words, the tongue raising (higher TH) is correlated with a decrease in F1 (Comivi Alowonou et al., 2019; Kent & Vorperian, 2018), which has been reported for both EP (Albuquerque et al., 2014) and other languages (Linville, 2001; Schötz, 2006; Torre III &

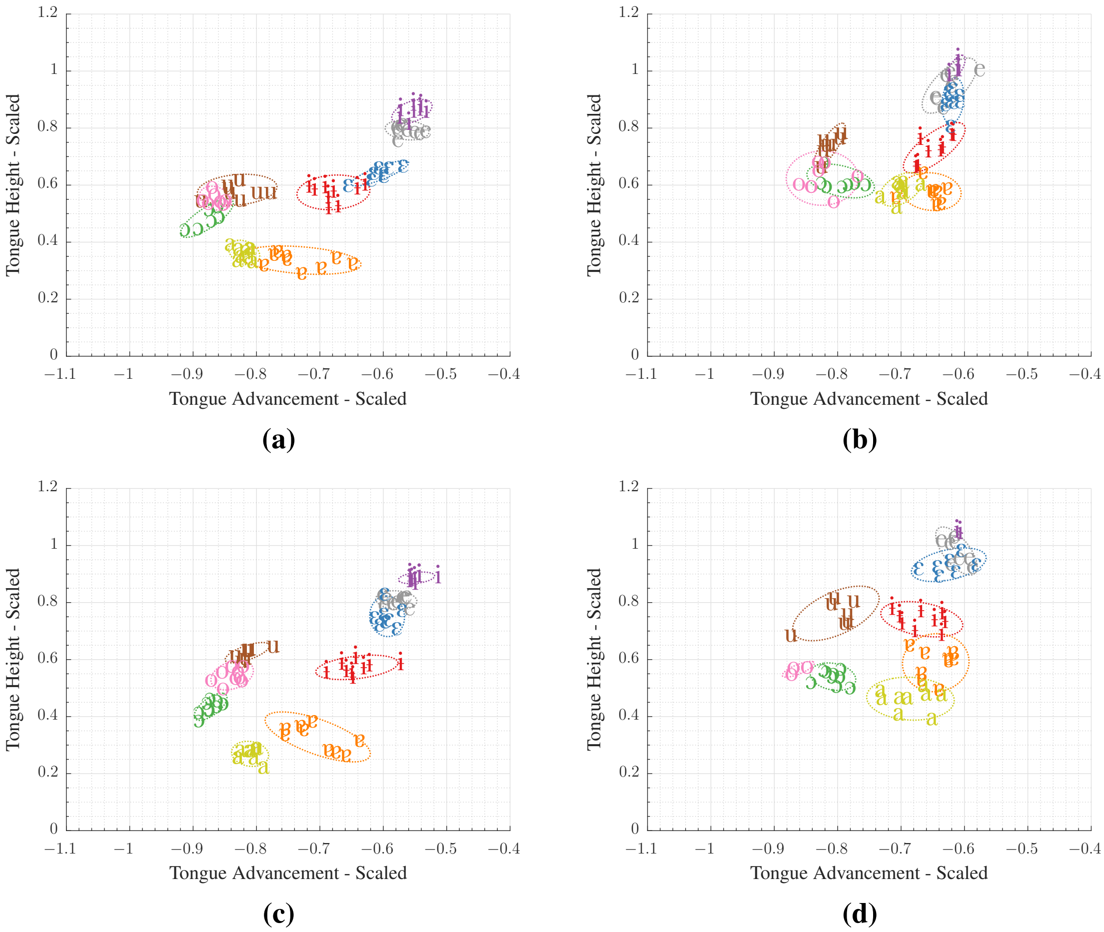


Figure 11.8: Vowel articulatory cluster size of all EP oral vowels of two speakers in pVCv sequences (top) (**a**) (**b**) and in isolated context (bottom) (**c**) (**d**). Left Side: young female (Y2); Right side: old female (O3). (adapted from Albuquerque, Valente, et al. (2022))

Barlow, 2009; P. J. Watson & Munson, 2007; Xue & Hao, 2003) with aging. However, Pellegrini et al. (2013) reported higher F1 values for vowels produced by older EP females.

Regarding TA, the results indicate a tendency to tongue advancement in the older group, that is correlated with an increase in F2 (Comivi Alowonou et al., 2019; Kent & Vorperian, 2018). The tongue advancement observed for older females is not in line with the previous acoustic data (Section 7), where a tendency for F2 decrease with aging occurred. Yet, some studies have reported vowel-specific changes with age and gender (Eichhorn et al., 2018; Pellegrini et al., 2013; Rastatter & Jacques, 1990; Schötz, 2006; Torre III & Barlow, 2009). For EP, Pellegrini et al. (2013) observed a trend for F2 decrease in back vowels and to increase in front vowels. Nonetheless, Sonies, Baum, and Shawker (1984) reported a reduction in tongue retraction on the vowel [a] in older adults, and suggested that older speakers tend to hold the tongue in a more anterior position in the mouth than younger speakers. The authors (Sonies, Baum, & Shawker, 1984) attributed this reduction in tongue retraction to a normal compensatory tongue motion, apparently due to alterations in the function of the suspensory muscles of the tongue (Sonies, Stone, & Shawker, 1984).

Beyond the observed displacement of the articulatory vowel space with age (i.e., more advanced) (see Figure 11.7), the vowels' space tends to be more reduced in younger than in older females. For EP, Pellegrini et al. (2013) also reported an acoustic space more reduced for younger females. Although formant frequencies can be affected by other articulatory adjustments than the tongue movements, it would be interesting to study acoustic-articulatory relationship of the data. Furthermore, Xue and Hao (2003) suggested that the lowering of formants with age might be due to the lengthening of the vocal tract with aging, but in the present study a tongue raising and a tongue advancement were observed for the older females.

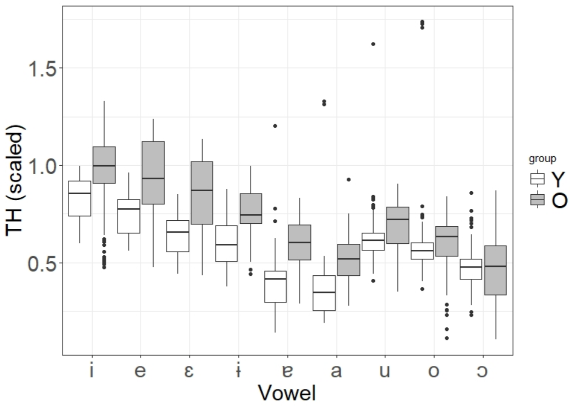
Additionally, the articulatory measures show great variability, as it can be observed by the different format of the individual articulatory vowel spaces, and the vowel articulatory clusters. The comparison of the amount of variability within and across speakers in both age groups needs to take into account the fact that older speakers present less data, which does not allow to draw solid conclusions. In addition, there does not appear to be a direct relationship between speaker size (BMI) and the length of the tongue contour or the size of the acoustic triangle obtained.

Concerning the vowel type, the results indicate a significant difference in tongue height position between the EP oral vowels examined, mainly between front and back vowels. As expected, for each block of vowels the TH is higher for close vowels and lower for open vowels. An interesting observation is that for young females the central vowels [ɐ] and [a] show lower values and similar TH among each other, while for older females the vowel [ɐ] presents higher TH than vowel [a]. Regarding TA, front vowels tend to present a more advanced higher point of the tongue contour.

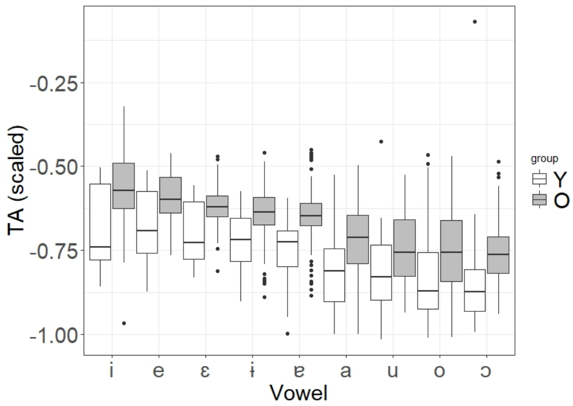
Regarding the vowel context, as in the pilot study (Albuquerque, Valente, Barros, et al., 2021), the articulatory space tends to be smaller when vowels occur in pVCv sequences comparing with isolated vowels, mainly for younger females. The vowel articulatory space reduction observed for vowels in pVCv sequences in comparison with isolated vowels might be related with the tendency to hyperarticulate isolated vowels. This type of effect was also evident between vowels in clear speech versus in conversational speech (Song, 2017), or in long vowels versus short vowels (W.-S. Lee, 2016), for other languages.

While for young females a reduction in the articulatory vowel space between contexts was observed, for older females the vowel space tends to be more advanced for isolated vowels. This different pattern between vowel contexts with age might be related to specific articulatory adjustments of the older females.

In this study the extraction of the tongue contours (using a U-Net adapted from Ronneberger et al. (2015)), and the determination of the highest position of the tongue body for young and old Portuguese females were done automatically. Figure 11.9 shows the TA and TH scaled measures obtained of all tongue contours without any manual verification. Compared with the data obtained after manual inspection of the tongue contours (see Figure 11.5,) similar tendencies for each vowel by age group are observed, even though a higher amount of outliers occur. In a broader perspective, this automatic method may enable an unsupervised extraction and analysis of these tongue measures, making possible the analysis of large amounts of US data for young and old speakers. Notwithstanding, more methodological work is necessary to improve the automatic tongue tracing to deal with extraction errors, and the accuracy of this method still needs to be measured, in order to avoid the manual revision of the tongue contours, as this is a time-consuming task.



(a)



(b)

Figure 11.9: Boxplots of tongue height (TH) (**a**) and tongue advancement (TA) (**b**) scaled values by vowel and age group of all tongue contours automatically obtained. (adapted from Albuquerque, Valente, et al. (2022))

As the dataset is small, it cannot be excluded that speaker-specific strategies interact with group behavior (Mücke, Thies, Mertens, & Hermes, 2020). Also, the US images could be affected by many factors, such as the anatomical differences (e.g., facial profile, thickness of the fat tissue around the neck), or the pressure or position of the transducer between speakers (Chantaramanee et al., 2019; Strycharczuk & Scobbie, 2017). Further, the noisy nature of the images make the segmentation demanding, and several unclear images had to be excluded from the study.

Moreover, the analysis of the TH and its horizontal location (TA) can be inaccurate when the highest point of the tongue is not located at the narrowest point of the tongue–palate constriction (Georgeton et al., 2016). However, as the palate traces were largely unreliable for the majority of the speakers, it is not possible to implement a measure of constriction degree for the current data. Thus, reducing the lingual configuration to a single point is a convenient methodological solution, but it is far from adequate in giving a comprehensive description of the lingual articulation (Georgeton et al., 2016).

As raw distance measures between vowels are sensitive to anatomical differences between speakers, this study proposes a new normalization approach that is independent of the proportion of the tongue imaged for each speaker. However, this normalization method requires further validation.

This page intentionally contains only this sentence.

PART IV

DISCUSSION AND CONCLUSIONS

This page intentionally contains only this sentence.

DISCUSSION AND CONCLUSIONS

In this chapter, an analysis and discussion of the main results obtained from the five studies presented above is carried out. A summary of the main findings of each study is included in the Section 12.1, followed by an overall discussion (Section 12.2) and the study limitations (Section 12.3). Finally, the future work is presented in Section 12.4.

12.1 Summary of the main findings

The studies comprising this thesis on age effects in EP speech include both segmental and suprasegmental levels. These studies were aimed at contributing to a comprehensive analysis of age and gender effects in EP speech using acoustic and articulatory data.

In order to achieve the main goal of this thesis, the need to create speech databases (with acoustic and articulatory data) emerged, due to the lack of EP speech data available for analysis. Therefore, the development of speech databases, with acoustic and articulatory data of healthy adults using standardized recording procedures, were crucial.

In this vein, the present work provided two speech databases for EP: i) one acoustic database containing all EP oral vowels produced in similar context (reading speech), and also semi-spontaneous speech (image description) collected from a large sample of adults of the central region of Portugal, between the ages 35 and 97; ii) and another one with articulatory data (US images) for all EP oral vowels produced in similar contexts (pseudowords and isolated) collected from young ([21-35]) and older ([55-73]) adults.

These databases were analyzed in five studies, where various aspects of the aging speech were explored through different methods, such as static acoustic features, dynamic vowel formants and articulatory data. This section summarizes the main findings and contributions of this work, and reviews the conclusions drawn from the studies presented in the previous chapters.

12.1.1 Age-related acoustic changes in EP oral vowels

The first study (Chapter 7) was developed in order to analyze the age effects on duration, F0 and formant frequencies (i.e., F1, F2 and F3) for all EP oral vowels produced by 112 healthy speakers of both genders between the ages 35 and 97. A selection of vowel space metrics (VSA, F1RR, F2RR, VAI and FCR) was also analyzed. In addition, the analyses considered age as a continuous variable (to avoid the arguable division into age groups).

In a first analysis, the results of this study provide a base of information to establish vowel acoustics normal patterns of aging among Portuguese adults. Summing up, the data have shown that, with aging: the duration of vowels increased in both genders; vowels' F0 tended to be closer between genders (i.e., vowels' F0 increased for males and slightly decreased for females); there was a general tendency for formant frequencies to decrease for females; F1 and F2 space underwent a reduction in males, confirmed by the vowel space acoustic indexes F1RR, F2RR, VAI and FCR; and F3 was generally unchanged, in both genders. The statistical analyses have also shown which vowels are more affected by aging: 1) the unstressed oral vowels F0, mainly for females (i.e., the F0 decreases significantly); 2) formants of vowels [u] and [ɔ] for females (i.e., F1 and F2 decrease); 3) vowels [i], [u], [a] and [e] males' formants (with a consequent vowel space centralization).

The results of the present study support some previously reported findings of the age-related changes on vowel acoustics, but also include novel results that deepen the understanding of the aging speech, mainly for the EP. First, some acoustic parameters (e.g., F0 and formant frequencies) did not yield as much variation with aging as reported for other languages (Eichhorn et al., 2018; Kyriaki, 2021; Torre III & Barlow, 2009; Vipperla et al., 2010). Despite non-significant, vowels' F0 decreased for females of about 25 Hz between the ages 35 and 100, and increased around 10 Hz for males between the same ages, in line to what was observed previously by Pellegrini et al. (2013) for EP.

Second, in contrast to Fletcher et al. (2015), which suggested that the vowel acoustic space remains unchanged with age due to the increase of the vowel duration, in the current study males revealed a significant reduction of the vowel acoustic space with age despite a significant increase of the vowels duration. That is, as reported by Gahl and Baayen (2019), changes in vowel duration do not inevitably result in articulatory movement alterations or vowel space contraction or expansion.

Finally, beyond the general properties of Portuguese vowels (Escudero et al., 2009) observed in this study, the older adults showed a lower front-back distinction compared with the younger ones (i.e., the differences of F0 and F1 between front and back vowels decreased with age).

12.1.2 Age-related effect on dynamic characteristics of vowel formants

The second study (Chapter 8) intended to understand the usefulness of dynamic features for classification tasks and characterization of vowel acoustics.

The performance of several age and vowel classification models was investigated with a dataset of oral vowels from 112 EP adult speakers. Features consisted of the first 3 DCT coefficients (C0 to C2) and a set of five representative types of classifiers was used.

Foremost, the accuracy results of age classification experiments showed improvement with the addition of dynamic features for the several age divisions considered (i.e., 2 age groups (non older/older and younger/older), three age groups ([35-45][46-69][70-97]) and four age groups ([35-49][50-64][65-79][80-97])). This improvement was also noticeable in vowel classification but of lower magnitude. Beyond that, adding vowels' duration had a positive impact not only on vowel classification but also on age classification, which is in accordance with the age-related vowel duration increase observed in the data and also reported for other languages (Benjamin, 1982; D'Alessandro & Fougeron, 2018; Fletcher et al., 2015; Kyriaki, 2021; Mertens et al., 2020; Schötz, 2006; Tykalova et al., 2020). Furthermore, considering the improvements related to chance level, the age division in three groups (i.e., [35-45], [46-69] and [70-97]) seems to be the one that fits better the lifespan speech alterations, since this age division addresses the menopause stage for females, where the younger group [35-45] consists only by women in the reproductive phase.

Statistical tests established a connection of the dynamic features with age: (1) vowel formant dynamics, particularly C1 of F1, was affected by age (i.e., higher trajectory slope); (2) some gender differences were found in C2 of F1 and F2 (i.e., in the curvature of the vowel formants), similarly to what was reported for the static information of vowel formants (i.e., C0). The fact that the F1 trajectory slope increases with aging for both genders may indicate that older adults present vowels with greatest amount of spectral change (i.e., "more exaggerated articulation") (Jacewicz et al., 2011a).

The results of this study tend to support the hypothesis that dynamic features of vowels carry important information about the speaker (McDougall & Nolan, 2007) and seem to be affected by age. Moreover, dynamic measurements of the vowels' formant would be highly desirable to provide a more complete view of vowel characteristics, and to avoid a necessarily arbitrary choice of selecting a specific time point where the measurements are taken.

12.1.3 The effects of age on suprasegmental acoustic parameters

The third study (Chapter 9) explored the age effects at the suprasegmental level in EP semi-spontaneous speech (obtained through a picture description task) produced by 112 adults, aged between 35 and 97.

Fundamentally, as age progresses, male participants tended to talk less (i.e., the total speech duration decreased), with a decrease in variability of speech interval duration and with higher pause time (i.e., more pause time and longer pause duration), which are in line with the literature for other languages (Dimitrova et al., 2018; Fougeron et al., 2021; Hartman & Danhauer, 1976; Massimo & Elisa, 2014; Pellegrini et al., 2013; Steffens, 2011). Female speakers presented a significant increase in speech and articulation rate and also a HNR decrease, which means that women speak faster and with a lower voice quality. With aging, more additive noise in the voiced signal is expected (Dehqan et al., 2013; Ferrand, 2002; Lortie et al., 2015; Xue & Deliyski, 2001). In contrast, the faster articulation and speech rate in older females is not in agreement with the general trend observed in previous studies (Fougeron et al., 2021; Goy et al., 2016; Hazan et al., 2018; Hermes et al., 2020; Pellegrini et al., 2013; Tykalova et al., 2020; Volín et al., 2017).

As in vowels' F0, but less markedly, speaking F0 also tended to decrease in women and to slightly increase in men with aging. However, it is possible that speaking F0 presents a non-linear variation with age, mainly in males, which should be further explored.

The results of this study provide, essentially, a starting point to establish the normal patterns of rhythm and intonation in semi-spontaneous speech across age among adult Portuguese native speakers. In general, these findings are in line with previous research suggesting that suprasegmental characteristics of speech change with age, with some gender differences.

12.1.4 Association between acoustic speech features and non-severe levels of mood symptoms across lifespan

The fourth study (Chapter 10) reported the association between non-severe symptoms of anxiety/depression and segmental and suprasegmental acoustic features. Thus, this study analyzed the impact of anxiety/depression symptoms on acoustic features extracted by a self-assessment of mood (using the HADS scale) in 112 individuals aged 35-97.

The number of participants of the present study with presence of anxiety or depression symptoms was low (>7: 26.8% and 10.7%, respectively). For these individuals, mainly constituted by adults with non-severe mood symptoms (HADS-A: 5.4 ± 2.9 and HADS-D: 4.2 ± 2.7, respectively), an increase in depressive symptoms is associated with higher vowel duration, increase of total pause duration and less total speech duration in the univariable model. Adjusting the linear model for age and gender revealed that age affects the depressive symptoms. Only the total speech duration decrease in the adjusted model, along with age, maintaining the significant association with depression symptoms. Contrariwise, an increase of the anxiety symptoms did not present significant relationships associated with the acoustic parameters studied.

Additionally, the fact that the older participants tended to have more depressive symptoms is in accordance with epidemiological studies developed in low-income countries, but not in Western countries, where the rate of depression decreases with aging (Bromet et al., 2005; Kessler et al., 2010). An increase of depression in older adults could reflect the low levels of cognitive and physical health observed in Portugal compared to other European countries (Balabanova & McKee, 2002; Bischoff-Ferrari, 2019; Bobak et al., 2004; Bromet et al., 2005; Schietzel et al., 2022).

In summary, some acoustic parameters are sensitive to depression symptoms and age, even among individuals without severe symptom levels.

12.1.5 Articulatory analysis of the age effect on EP vowel production

In the fifth study (Chapter 11), the age-related articulatory changes were investigated through the analysis of the US tongue contours of 9 EP oral vowels (in isolated context and in pseudoword context) produced by 10 female speakers of two different age groups (young and old).

First, an improved automatic method for tongue contour tracing allowed expanding the number of considered speakers and supported a novel normalization procedure for inter-speaker comparison. Therefore, from the tongue contours automatically segmented from the US images and manually revised, the parameters TH and TA were extracted.

In light of this new articulatory data, it can be concluded that the tongue tends to be higher and more advanced with aging for almost all vowels, meaning that the vowel articulatory space tends to be higher, advanced, and bigger in older females. The US study of Sonies, Baum, and Shawker (1984) also suggested that older speakers tend to hold the tongue in a more anterior position in the mouth than younger speakers.

Concerning the vowel type, beyond the general properties of vowels observed in this study (i.e., differences in TH and TA based on the vowel position), it should be noted that for young females, the central vowels [ɐ] and [a] show lower values and similar TH among each other, while for older females, the vowel [ɐ] presents higher TH than vowel [a]. This fact may indicate some change in progress, i.e., diachronic or age-related changes.

Regarding the context, the vowel space tends to be more advanced for isolated vowels comparing with vowels produced in disyllabic sequences for older females, while younger females tend to present a sharp reduction in the articulatory vowel space in disyllabic sequences. This different pattern between vowel contexts with aging might be related to specific articulatory adjustments of the older females.

Globally, these results contribute to an accurate articulatory description of EP oral vowels, and also provide valuable insights about the vowel articulatory normal patterns of aging among female Portuguese adults. In addition, the measures TH and TA extracted of the US tongue data tend to confirm the usefulness of these metrics in characterizing and differentiating, at least, the corner vowels across both speaker groups.

12.2 Overall discussion

The current thesis presents both novel and supportive findings regarding age-related changes on speech production. Globally, the present results show that, acoustically, the aging speech of EP healthy adults is characterized by: 1) longer vowels (in both genders); 2) a tendency for F0 to decrease in women and slightly increase in men (more pronounced in vowels' F0 than in speaking F0); 3) lower vowel formant frequencies in females, especially F1 and F2; 4) a significant reduction of the vowel acoustic space in men; 5) vowels with higher trajectory slope of F1 (in both genders); 6) shorter descriptions with higher pause time for males; 7) faster speech and articulation rate for females (in semi-spontaneous speech); and 8) a possible decrease of the voice quality for females (i.e., lower HNR in semi-spontaneous speech).

First of all, these results corroborated that acoustic characteristics of speech change with age and present different patterns between genders. Only the increase of the vowels duration and the rise of the amount of spectral change for F1 are similar for both genders with aging. Additionally, the F3 of vowels remains unchanged with age in both genders, which does not support the idea of vocal tract lengthening with aging. Therefore, the different pattern of formant frequencies variation with age might be related to gender and/or specific articulatory adjustments of the older speakers during speech. Beyond the positional vowel changes, the age-related changes observed on formant dynamics (i.e., in F1C1) also confirm this statement. The results also suggested that vowel duration does not inevitably result in vowel space contraction or expansion (Gahl & Baayen, 2019).

Furthermore, as the literature have suggested that the segment duration depends on the speech rate (i.e., a faster speech rate results in a decrease of the segments duration) (Linville, 2001; Schötz, 2006; Smith et al., 1987), the increase of speech and articulation rate with age observed for females in semi-spontaneous speech is not in accordance with the increase of vowel duration. Nonetheless, the different type of speech corpus used to measure the vowel duration and the speech/articulation rate might influence these results. Thus, as semi-spontaneous speech requires longer utterances, older females might increase the articulation rate to maintain the physiological support needed to successfully produce the utterance (Linville, 2001). In contrast, older males show more pause time and longer pause duration that could also represent an aggravation of the lung-function (Schötz, 2006; Steffens, 2011). However, longer pauses also give more time for speakers to solve difficulties of speech planning or articulation (Bóna, 2014).

Despite non-significant, speaking F0 and vowels' F0 decrease for females of about 20-25 Hz between the ages 35 and 100, and increase around 5-10 Hz for males between the same ages. The lack of significant age effects for vowels' F0 and speaking F0 is surprising, particularly for women, as significant changes in these parameters have been observed in several studies of aging speech (Eichhorn et al., 2018; Mautner, 2011; Torre III & Barlow, 2009; Tykalova et al., 2020; Vipperla et al., 2010). However, lifespan studies of F0 could lead to different results depending on which age ranges are studied or the type of speech corpus used (Fougeron et al., 2021; Fuchs et al., 2021; Stathopoulos et al., 2011). As F0 depends on the physiological characteristics of the vocal folds and the control of the larynx musculature (Baken, 2005; Lortie et al., 2015; Makiyama & Hirano, 2017), the absence of significant changes observed in mean F0 suggests that the vibration pattern of the vocal folds did not deteriorate with aging (at least in the age range considered here), or did not deteriorate enough to have a functional impact. In addition, Lortie et al. (2015) also indicated that connected speech was less vulnerable to age than sustained vowels.

In relation to the association between the acoustic parameters with depressive and anxiety symptoms, non-severe depression symptoms can be related to the change of some acoustic parameters and age, specifically with the decrease of the total speech duration. That is, older adults tended to present more depressive symptoms that could impact the amount of speech produced and the communication process.

As the articulatory adjustments underlying the acoustic findings have been poorly explored, US data of the tongue contours were analyzed in this study. The current findings indicate that the vowel articulatory space tends to be higher, advanced, and bigger in older females. The tongue raising (higher TH) is correlated with a decrease in F1 (Comivi Alowonou et al., 2019; Kent & Vorperian, 2018), which has been reported for different languages with aging (Albuquerque et al., 2014; Linville, 2001; Schötz, 2006; Torre III & Barlow, 2009; P. J. Watson & Munson, 2007; Xue & Hao, 2003), and it is also in line with the results obtained in the acoustic study based on static information of vowel formants. Contrariwise, the tongue advancement, which is correlated with an increase in F2 (Comivi Alowonou et al., 2019; Kent & Vorperian, 2018), observed for older females, is not in line with the data obtained in the vowel acoustic study, where vowels' F2 tend to slightly decrease with aging (only significant for F2[ɔ]). Although a decrease in vowels' F2 can be caused by a more posterior tongue position, it may also occur due to an increase in lip rounding (Kent & Rountrey, 2020; Titze, 1994; Wieling et al., 2016). Furthermore, an increase in lip rounding also affects the F1, decreasing it. Thus, although both acoustic and US data were shown to be useful for detecting age-related speech changes, these previous observations reinforce the need to also investigate the lip movement in studies of aging speech.

Beyond the observed displacement of the articulatory vowel space with age, the vowels' space tends to be more reduced in younger than in older females, which is in accordance with Pellegrini et al. (2013) that also reported a more reduced acoustic space for EP younger females. Nonetheless, this is the opposite of what was observed for men in the acoustic data, which may presuppose that older adults develop different strategies depending on gender, as argued before.

To sum up, the results of this thesis provide a starting point to establish the normal patterns of aging speech among (healthy) adult native Portuguese speakers. This study provides information on the effects of age on speech at both segmental and suprasegmental levels showing that, acoustically, the age-related speech changes are more gradual throughout adulthood than previously described (mainly for other languages), and it can also occur non-linearly. Nonetheless, as the same acoustic output can be achieved through different articulatory strategies and different speakers have distinct oral morphologies, speakers tend to differ more in their articulatory strategies than in their acoustic output (Tabain, 2013). Thus, older adults might develop different articulatory strategies to maintain the same acoustic output.

Nonetheless, an open and interesting question remains: whether the changes that have been observed in the data are the result of passive physiological changes to the vocal tract, or whether speech production is actively modified with increasing age, in order to perceptually compensate for the influences of the age-related decline on vowel quality.

The acoustic and articulatory changes resulting from the natural process of aging are an important basis to understand speech disorders associated with the health conditions that affect older individuals (e.g. hearing loss, dentofacial alterations, neurodegenerative diseases, stroke, cancer, or psychological distress) (Eichhorn et al., 2018; Kent & Vorperian, 2018). Wherefore, it is very important that voice clinicians are aware of such effects and take these into account in their intervention. That is, these reference data for EP, not only about acoustic and articulatory features of vowels, but also about suprasegmental characteristics of speech, are an important benchmark for clinical assessment and treatment of different speech disorders that are often age-related. Furthermore, the aging acoustic correlates reported in this thesis may further be helpful for the development of methods for both the automatic detection and the synthesis of speaker age. Thereby, improved age recognizers or classifiers may be achieved, as well as better and more natural-sounding synthesis of speaker age (Schötz, 2006).

As the population aging increases, the world is facing new challenges, and the number of older individuals seeking to maintain an effective vocal communication for as long as possible also increase (Caruso et al., 1995). In this vein, this work provides important information for the clinicians and also for the speech technologies, which could be useful to maintain the communication skills and the quality of life in the older ages (since the communication skills are basic avenues of socialization (Caruso et al., 1995)).

Summing up, the speaker age is a very complex characteristic of speech that leaves traces in all phonetic dimensions and its impact on the speech is influenced by numerous factors (Makiyama & Hirano, 2017; Schötz, 2006). Nonetheless, this thesis allowed to try out a number of acoustic and articulatory speech research methods to study the aging speech, which contributed to increasing knowledge on this topic, and might inspire some readers to continue investigating the age effects on speech.

12.3 Limitations

Despite the new contributions of this thesis to the research field, some limitations need to be acknowledged. Overall, even though cross-sectional data provided important information on the age-related changes on speech production, some drawbacks can be pointed out: 1) the data are based on a single-occasion observation, which makes their generalization to other time periods questionable (Pellegrino, 2019); 2) the methodological differences across studies might lead to variable results (Kent & Vorperian, 2018), namely due to the different age groups considered; 3) the aging effects where based on chronological age as proxy of physiological changes, but changes with age will also be affected by other factors, such as diachronic changes, lifestyle habits, educational and social environment (Fuchs et al., 2020; Makiyama & Hirano, 2017; Reubold & Harrington, 2015; Schötz, 2006); and 4) the speech data only corresponds to the central region of Portugal.

Regarding the acoustic database, and the subsequent studies, some aspects could affect the results, namely the type of speech samples used (i.e., no conversational data or sustained vowels); the differences in the recording environment (i.e., no soundproof room); the non-control of vowel duration; the automatic extraction procedures (i.e., not all vowel formants and labeled syllables were manually verified, but they were obtained using the same procedure for all speakers); the lack of balance regarding the educational level (i.e., the number of speakers with higher educational level decreases with age, and the majority of the older speakers needed support of the researcher to answer the questionnaires). Despite the fact that the impact of the speaker age on speech is influenced by numerous factors, such as physiological condition, occupations and lifestyle habits (Makiyama & Hirano, 2017; Schötz, 2006), these aspects were reported to characterize the speakers, but were not dealt with in this thesis. In addition to the age and gender, only the association of the anxiety and depression scores with the acoustic features were analyzed.

In relation to the articulatory database, the main limitations are methodological in nature and correspond to the US image quality (or the noisy nature of the images) that makes the segmentation demanding and could also difficult the accurate determination of articulatory measures for some vowels. That is, as reported before, the US images could be affected by the anatomical differences (e.g., facial profile, thickness of the fat tissue around the neck), or the pressure or position of the transducer between speakers, among other factors (Chantaramanee et al., 2019; Strycharczuk & Scobbie, 2015). Additionally, the palate traces revealed themselves largely unreliable for the majority of the speakers and the recording of the bite plane (in order to image the occlusal plane) was also challenging.

Specifically for the US results (Chapter 11) the following drawbacks can be pointed out: 1) the dataset used is small and restricted to females (i.e., the speaker-specific strategies might interact with group behavior and statistical results by age group would not be relevant); 2) there are no tongue contours available for some vowels produced by the older speakers; 3) the lingual configuration was reduced to a single point, which is far from adequate in giving a comprehensive description of the lingual articulation (Georgeton et al., 2016).

12.4 Future research directions

In the scope of the goals of this thesis, several questions remained unaddressed and deserve further research studies, arising many and diverse opportunities for future research. The ones that appear more promising are presented below.

First, this thesis reinforces the need for further investigation of acoustic and articulatory characteristics in aging speech, since this work is the starting point for a broader lifespan study, involving a large number of EP speakers, from infancy to old age, producing different types of speech corpora. In addition, the non-linear variation of F0 with age, and also of other speech parameters, should be further explored in future studies.

Despite the study of the age-related effect on dynamic characteristics of vowel formants adding important new knowledge, relevant for advancing lifespan phonetic studies, it deserves continuation in several lines, mainly: 1) investigate the effects of age on dynamic properties of vowel formant frequencies (F1, F2), applying generalized additive mixed models (GAMMs) to compare curvilinearly varying formant measurements along vowel segments; 2) identify which are the dynamic features with more impact on classification performance; 3) consider other features besides formants, such as cepstral coefficients; 4) explore new and larger datasets; and 5) extend the set of machine learning algorithms, including, as soon as dataset allows, deep learning methods.

Although the association between non-severe symptoms of anxiety/depression and segmental and suprasegmental acoustic features was observed, future research studies should analyze acoustic features extracted from other speech samples (e.g., text reading) in a group of individuals with a diagnose of anxiety and/or depression compared with a control group across lifespan.

Considering that the present study is a preliminary analysis of age articulatory changes, it would be essential to extend the study to the male gender and to analyze other measures such as the tongue root and the total vowel contour through smoothing spline ANOVA (SSANOVA) (Mielke, 2015; Song, 2017; Turgeon et al., 2017). Also, static and dynamic studies of the vowel tongue traces need to be addressed to investigate articulatory movement and velocity. Even though the analysis of data from more speakers could be desirable, it is important to be cautious with US data as the image quality can vary considerably among speakers, and the data fidelity is the most important when drawing conclusions on this research field (Dawson, 2020).

Additionally, as formant frequencies can be affected by other articulatory adjustments than the tongue movements, it would be interesting to study the acoustic–articulatory relationship of the data using not only the US images but also other advanced articulatory techniques (e.g., RTMRI) that allow to study other articulators, such as the lips.

Finally, longitudinal research studies that trace the acoustic features of the same speaker over a long period of time should be developed to complement the data obtained (study already in progress (Valente et al., 2021)).

To finish, further studies about the effects of typical aging on speech mechanism may benefit clinically, to provide a framework for typical and disordered speech production assessment, and guide intervention practices appropriately (Belmont, 2015). An improved model of speech production and a more defined aging trajectory would be advantageous to develop ASR systems suitable for older speech. Also, the plan for a *Decade of Healthy Ageing 2020–2030* (World Health Organization, 2020) identifies the need to “promote the development, production, availability and use of assistive and digital technologies and innovations that increase access to good-quality health and social services” by older people (World Health Organization, 2020, p.14). In this sense, as the older population is continuously growing and will comprise an important segment in the economy and society, it is important to produce scientific data to understand the normal aging and ensure the quality of life in the older ages.

This page intentionally contains only this sentence.

REFERENCES

Aalami, O. O., Fang, T. D., Song, H. M., & Nacamuli, R. P. (2003). Physiological features of aging persons. *Archives of Surgery*, *138*(10), 1068–1076. doi: 10.1001/ARCHSURG.138.10.1068

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

Adda-Decker, M., Boula De Mareüil, P., & Lamel, L. (1999). Pronunciation variants in French: Schwa & liaison. In *ICPhS99* (pp. 2239–2242). San Francisco.

Akgul, Y. S., Stone, C., & Maureen, K. (1999). Automatic extraction and tracking of contours. *IEEE Transactions on Medical Imaging*, *18*(10), 1035–1045.

Albuquerque, L., Oliveira, C., Teixeira, A., & Figueiredo, D. (2021). Eppur si muove: Formant dynamics is relevant for the study of Speech Aging Effects. In *14th BIOSTEC* (pp. 276–283). Online. doi: 10.5220/0010320902760283

Albuquerque, L., Oliveira, C., Teixeira, A., Sa-Couto, P., & Figueiredo, D. (2020). A comprehensive analysis of age and gender effects in European Portuguese oral vowels. *Journal of Voice*(in press). doi: 10.1016/J.JVOICE.2020.10.021

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

Albuquerque, L., Teixeira, A., Oliveira, C., & Figueiredo, D. (2022). Age and vowel classification improvement by the inclusion of vowel dynamic features. *International Journal of Speech Technology*, *25*, 1025–1040. doi: 10.1007/s10772-022-10012-x

Albuquerque, L., Valente, A. R., Barros, F., Teixeira, A., Silva, S., Martins, P., & Oliveira, C. (2021). The age effects on EP vowel production: An ultrasound pilot study. In *IberSPEECH 2021* (pp. 245–249). Valladolid, Spain: ISCA. doi: 10.21437/IberSPEECH.2021-52

Albuquerque, L., Valente, A. R., Barros, F., Teixeira, A., Silva, S., Martins, P., & Oliveira, C. (2022, jan). Exploring the age effects on European Portuguese vowel production: An ultrasound study. *Applied Sciences*, *12*(3), 1396. doi: 10.3390/app12031396

Albuquerque, L., Valente, A. R., Teixeira, A., Figueiredo, D., Sa-Couto, P., & Oliveira, C. (2021, apr). Association between acoustic speech features and non-severe levels of anxiety and depression symptoms across lifespan. *PLOS ONE*, *16*(4), e0248842. doi: 10.1371/journal.pone.0248842

Albuquerque, L., Valente, A. R., Teixeira, A., Oliveira, C., & Figueiredo, D. (2021). Age and gender effects in European Portuguese spontaneous speech. *Loquens*, *8*(1-2), e077. doi: 10.3989/LOQUENS.2021.077

Alghowinem, S., Goecke, R., Wagner, M., Epps, J., Breakspear, M., & Parker, G. (2012). From joyous to clinically depressed: Mood detection using spontaneous speech. In *25th International Florida Artificial Intelligence Research Society Conference.* Association for the Advancement of Artificial Intelligence (AAAI).

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

Alpert, M., Pouget, E. R., & Silva, R. R. (2001). Reflections of depression in acoustic measures of the patient's speech. *Journal of affective disorders*, *66*(1), 59–69.

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

Ambreen, S., Bashir, N., Tarar, S. A., & Kausar, R. (2019). Acoustic analysis of normal voice patterns in Pakistani adults. *Journal of Voice*, *33*(1), 124.e49–124.e58. doi: 10.1016/j.jvoice.2017.09.003

American Psychological Association. (2020). *APA Dictionary of Psychology.* Retrieved 2022-01-31, from `https://dictionary.apa.org/adulthood`

Arias-Vergara, T., Vásquez-Correa, J. C., & 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.

Articulate Assistant Ltd. (2014). *Articulate Assistant Advanced ultrasound module user manual.*

Articulate Instruments Ltd. (2008). *Ultrasound stabilisation headset users manual.* Edinburgh, UK: Articulate Instruments Ltd.

Articulate Instruments Ltd. (2010). *SyncBrightUp users manual.* Edinburgh, UK: Articulate Instruments Ltd.

Arviso, L. C., & Johns III, M. M. (2009). The aging voice. *Jornal of ENT Masterclass*, *2*(1), 44.

Audibert, N., Fougeron, C., Gendrot, C., & Adda-Decker, M. (2015). Duration- vs. style-dependent vowel variation: A multiparametric investigation. In *ICPhS'15* (pp. 5–9). Glasgow.

Baba, A., Yoshizawa, S., Yamada, M., Lee, A., & Shikano, K. (2004). Acoustic models of the elderly for large-vocabulary continuous speech recognition. *Electronics and Communications in Japan (Part II: Electronics)*, *87*(7), 49–57.

Baghban, K., Zarifian, T., Adibi, A., Shati, M., & Derakhshandeh, F. (2020). The quantitative ultrasound study of tongue shape and movement in normal Persian speaking children. *International Journal of Pediatric Otorhinolaryngology*, *134*, 110051. doi: 10.1016/j.ijporl.2020.110051

Baken, R. J. (2005). The aged voice: A new hypothesis. *Journal of Voice*, *19*(3), 317–325.

Baker, K. K., Ramig, L. O., Sapir, S., Luschei, E. S., & Smith, M. E. (2001). Control of vocal loudness in young and old adults. *J Speech Lang Hear Res*, *44*(2), 297–305. doi: 10.1044/1092-4388(2001/024)

Balabanova, D., & McKee, M. (2002). Access to health care in a system transition: the case of Bulgaria. *Int J Health Plann Mgmt*, *17*(4), 377–395. doi: 10.1002/hpm.687

Balcombe, N. R., & Sinclair, A. (2001). Ageing: definitions, mechanisms and the magnitude of the problem. *Best Practice & Research Clinical Gastroenterology*, *15*(6), 835–849. doi: 10.1053/BEGA.2001.0244

Ball, M. (2021). *Manual of clinical phonetics*. Routledge.

Banse, R., & Scherer, K. R. (1996). Acoustic profiles in vocal emotion expression. *Journal of Personality and Social Psychology*, *70*(3), 614–636. doi: 10.1037/0022-3514.70.3.614

Barbier, G., Perrier, P., Ménard, L., Payan, Y., Tiede, M., & Perkell, J. (2015). Speech planning in 4-year-old children versus adults: Acoustic and articulatory analyses. In *INTERSPEECH.*

Barbosa, P. A. (2006). *Incursões em torno do ritmo da fala (Incursions around the rhythm of speech).* Campinas: FAPESP/Pontes Editores.

Barbosa, P. A. (2010). Automatic duration-related salience detection in Brazilian Portuguese read and spontaneous speech. In *Speech Prosody.* Chicago.

Barbosa, P. A. (2013). Semi-automatic and automatic tools for generating prosodic descriptors for prosody research. In *TRASP* (Vol. 13, pp. 86–89). Aix-en-Provence.

Barbosa, P. A. (2019). *Prosódia (Prosody)* (Vol. 2). São Paulo: Parábola.

Barbosa, P. A., & Madureira, S. (2015). *Manual de fonética acústica experimental: aplicações a dados do português (Handbook of experimental acoustic phonetics: Applications to portuguese data).* Cortez Editora.

Barnes, S. F. (2011a). Fourth age – The final years of adulthood. In *San Diego State University Interwork Institute.* doi: 10.1159/000067946

Barnes, S. F. (2011b). Third age – The golden years of adulthood. In *San Diego State University Interwork Institute.* doi: 10.1037/a0021285

Barros, F., Silva, S., Albuquerque, L., Valente, A. R., Teixeira, A., Martins, P., & Oliveira, C. (2020). Towards the use of ultrasonography to study aging effects in vowel production [Poster]. In *12th ISSP.* Online.

Barros, F., Valente, A. R., Albuquerque, L., Silva, S., Teixeira, A., & Oliveira, C. (2020). Contributions to a quantitative unsupervised processing and analysis of tongue in ultrasound images. *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, *12132 LNCS*, 170–181. doi: 10.1007/978-3-030-50516-5_15

Bässler, R. (1987). Histopathology of different types of atrophy of the human tongue. *Path. Res. Pract.*, *182*(1), 87–97. doi: 10.1016/S0344-0338(87)80147-4

Belmont, A. J. (2015). *Anticipatory coarticulation and stability of speech in typically fluent speakers and people who stutter across the lifespan: An ultrasound study* (Master's thesis). University of South Florida.

Benjamin, B. J. (1981). Frequency Variability in the Aged Voice. *Journal of Gerontology*, *36*(6), 722–726. doi: 10.1093/geronj/36.6.722

Benjamin, B. J. (1982). Phonological performance in gerontological speech. *Journal of Psycholinguistic Research*, *11*(2), 159–167. doi: 10.1007/bf01068218

Bennabi, D., Vandel, P., Papaxanthis, C., Pozzo, T., & Haffen, E. (2013). Psychomotor retardation in depression: a systematic review of diagnostic, pathophysiologic, and therapeutic implications. *BioMed Research International*, 1–18. doi: 10.1155/2013/158746

Berg, M., Fuchs, M., Wirkner, K., Loeffler, M., Engel, C., & Berger, T. (2017). The speaking voice in the general population: Normative data and associations to sociodemographic and lifestyle factors. *Journal of Voice*, *31*(2), 257.e13–257.e24. doi: 10.1016/j.jvoice.2016.06.001

Bilodeau-Mercure, M., & Tremblay, P. (2016). Age differences in sequential speech production: Articulatory and physiological factors. *Journal of the American Geriatrics Society*, *64*(11).

Birren, J. E., & Schaie, K. W. (2001). *Handbook of the psychology of aging* (5a ed.). San Diego: Academic Press.

Bischoff-Ferrari, H. A. c. (2019). *DO-HEALTH.* Retrieved from `http://do-health.eu/`

wordpress/

Bjelland, I., Dahl, A. A., Haug, T. T., & Neckelmann, D. (2002). The validity of the Hospital Anxiety and Depression Scale: an updated literature review. *Journal of psychosomatic research*, *52*(2), 69–77.

Blayer, I. M. F. (2002). On vocalic* variations in Portuguese dialects. *Signum: Estudos da Linguagem*, *5*(1), 77–106.

Bobak, M., Room, R., Pikhart, H., Kubinova, R., Malyutina, S., Pajak, A., ... Marmot, M. (2004). Contribution of drinking patterns to differences in rates of alcohol related problems between three urban populations. *Journal of Epidemiology and Community Health*, *58*(3), 238 LP – 242. doi: 10.1136/jech.2003.011825

Boersma, P., & Weenink, D. (2012). *Praat: doing phonetics by computer* (5.3.23 ed.). University of Amsterdam. Retrieved from http://www.praat.org/

Boléo, M. D. P., & Silva, M. H. S. (1961). O mapa dos dialetos e falares de Portugal continental (The map of dialects and speeches of mainland Portugal). In *Boletim de filologia* (pp. 85–112). Lisbon, Portugal: Centro de Estudos Filológicos.

Bóna, J. (2014). Temporal characteristics of speech: The effect of age and speech style. *J. Acoust. Soc. Am.*, *136*(2), EL116–EL121. doi: 10.1121/1.4885482

Boone, D. R., McFarlane, S. C., Berg, S. L. V., & Zraick, R. I. (2010). *The voice and voice therapy* (8a ed.). Upper Saddle River, N.J.: Pearson Education.

Bourbon, A., & Hermes, A. (2020). Have a break: Aging effects on sentence production and structuring in French. In *12th International Seminar on Speech Production (ISSP).*

Branco, A., Fabro, A. T., Gonçalves, T. M., & Martins, R. H. G. (2015, feb). Alterations in extracellular matrix composition in the aging larynx. *Otolaryngology - Head and Neck Surgery (United States)*, *152*(2), 302–307. doi: 10.1177/0194599814562727

Branco, A., Rodrigues, S. A., Fabro, A. T., Fonseca-Alves, C. E., & Martins, R. H. G. (2014, oct). Hyaluronic acid behavior in the lamina propria of the larynx with advancing age. *Otolaryngology - Head and Neck Surgery*, *151*(4), 652–656. doi: 10.1177/0194599814544673

Breiman, L. (2001). Random forests. *Machine Learning*, *45*(1), 5–32. doi: 10.1023/A:1010933404324

Breznitz, Z. (1992). Verbal indicators of depression. *The Journal of General Psychology*, *119*(4), 351–363. doi: 10.1080/00221309.1992.9921178

Brissos, F. (2011). *Linguagem do Sueste da Beira no tempo e no espaço (Language of southeastern Beira in time and space)* (Doctoral dissertation). Faculdade de Letras da Universidade de Lisboa.

Brissos, F. (2018). Proposta de reformulação da classificação dialetal do noroeste português (Proposal to reformulate the dialectal classification of Portuguese northwest). *Estudos de Linguistica Galega*, 193–208.

Brissos, F. (2020). Sistemas vocálicos tónicos portugueses do norte e do centro-sul: descrição acústica e tendências dialetais (Portuguese stressed vowel systems from the north and center-south: acoustic description and dialectal tendencies). *Zeitschrift für romanische Philologie*, *136*(1), 161–227. doi: 10.1515/zrp-2020-0007

Brissos, F., & Saramago, J. (2014). O problema da diversidade dialectal do Centro-Sul português: informação perceptiva versus informação acústica (The problem of dialectal diversity in the

Portuguese Center-South: perceptual information versus acoustic information). *Estudos de Linguistica Galega*, *6*, 53–80.

Bromet, E. J., Gluzman, S. F., Paniotto, V. I., Webb, C. P. M., Tintle, N. L., Zakhozha, V., ... Schwartz, J. E. (2005). Epidemiology of psychiatric and alcohol disorders in Ukraine. *Soc Psychiatry Psychiatr Epidemiol*, *40*(9), 681–690. doi: 10.1007/s00127-005-0927-9

Broskey, N. T., Marlatt, K. L., Most, J., Erickson, M. L., Irving, B. A., & Redman, L. M. (2019). The panacea of human aging: Calorie restriction versus exercise. *Exerc Sport Sci Rev.*, *47*(3), 169–175. doi: 10.1249/JES.0000000000000193

Brückl, M., & Sendlmeier, W. (2003). Aging female voices: An acoustic and perceptive analysis. In *Voqual'03* (pp. 163–168). Geneva, Switzerland.

Bruzzi, C., Salsi, D., Minghetti, D., Negri, M., Casolino, D., & Sessa, M. (2017). Presbyphonia. *Acta Biomedica*, *88*(1), 6–10. doi: 10.23750/ABM.V88I1.5266

Cannizzaro, M., Harel, B., Reilly, N., Chappell, P., & Snyder, P. J. (2004). Voice acoustical measurement of the severity of major depression. *Brain and cognition*, *56*(1), 30–35.

Carl, M. (2018). *Vowel production in Down syndrome: An ultrasound study* (Doctoral dissertation). City University of New York.

Caruso, A. J., Mueller, P. B., & Shadden, B. B. (1995). Effects of aging on speech and voice. *Physical & Occupational Therapy in Geiiatrics*, *13*(1-2), 63–79. doi: 10.1080/J148V13N01_04

Caverlé, M. W. J., & Vogel, A. P. (2020). Stability, reliability, and sensitivity of acoustic measures of vowel space: A comparison of vowel space area, formant centralization ratio, and vowel articulation index. *J. Acoust. Soc. Am.*, *148*(3), 1436–1444. doi: 10.1121/10.0001931

Chalise, H. N. (2019). Aging: basic concept. *Am J Biomed Sci & Res.*, *1*(1), 8–10. doi: 10.34297/AJBSR.2019.01.000503

Chantaramanee, A., Tohara, H., Nakagawa, K., Hara, K., Nakane, A., Yamaguchi, K., ... Minakuchi, S. (2019). Association between echo intensity of the tongue and its thickness and function in elderly subjects. *Journal of oral rehabilitation*.

Chittaragi, N. B., & Koolagudi, S. G. (2019). Acoustic-phonetic feature based Kannada dialect identification from vowel sounds. *International Journal of Speech Technology*, *22*(3), 1099–1113. doi: 10.1007/s10772-019-09646-1

Cintra, L. F. L. (1971). Nova proposta de classificação dos dialetos Galego-Portugueses (New proposal for the classification of Galician-Portuguese dialects). *Boletim de Filologia*, *22*, 81–116.

Comivi Alowonou, K., Wei, J., Lu, W., Liu, Z., Honda, K., & Dang, J. (2019). Acoustic and articulatory study of Ewe vowels: A comparative study of male and female. In *INTERSPEECH* (pp. 1776–1780). Graz, Austria: ISCA. doi: 10.21437/Interspeech.2019-2196

Costa, F. (2004). Intrinsic prosodic properties of stressed vowels in European Portuguese. In *Speech prosody* (pp. 53–56). Nara, Japan: ISCA.

Costa, J. P., Vitorino, R., Silva, G. M., Vogel, C., Duarte, A. C., & Rocha-Santos, T. (2016). A synopsis on aging - Theories, mechanisms and future prospects. *Ageing Research Reviews*, *29*, 90–112. doi: 10.1016/J.ARR.2016.06.005

Cox, V. O., & Selent, M. (2015). Acoustic and respiratory measures as a function of age in the male voice. *J Phonet and Audiol*, *1*(105), 2.

Crow, H. C., & Ship, J. A. (1996). Tongue strength and endurance in different aged individuals. *Journal of Gerontology: MEDICAL SCIENCES*, *51A*(5), M247–M250.

Crystal, D. (1969). *Prosodic systems and intonation in English* (Vol. 1). CUP Archive.

Cuesta, P. V., & Luz, M. A. M. D. (1980). *Gramática da língua portuguesa (Grammar of the portuguese language)*. Lisboa: Edições 70.

Cummins, N., Epps, J., & Ambikairajah, E. (2013). Spectro-temporal analysis of speech affected by depression and psychomotor retardation. In *2013 IEEE International Conference on Acoustics, Speech and Signal Processing* (pp. 7542–7546). IEEE.

Cummins, N., Epps, J., Breakspear, M., & Goecke, R. (2011). An investigation of depressed speech detection: Features and normalization. In *INTERSPEECH* (pp. 2997–3000).

Cummins, N., Scherer, S., Krajewski, J., Schnieder, S., Epps, J., & Quatieri, T. F. (2015). A review of depression and suicide risk assessment using speech analysis. *Speech Communication*, *71*, 10–49.

Cummins, N., Sethu, V., Epps, J., Schnieder, S., & Krajewski, J. (2015). Analysis of acoustic space variability in speech affected by depression. *Speech Communication*, *75*, 27–49.

Cunha, C., Silva, S., Teixeira, A., Almeida, N., Frahm, J., & Joseph, A. (2020). First results of an RT-MRI study on the influence of context in the synchronization of European Portuguese nasal vowels. In *12th ISSP (Poster).* Online.

Cunha, C., Silva, S., Teixeira, A., Oliveira, C., Martins, P., Joseph, A., & Frahm, J. (2019). On the role of oral configurations in European Portuguese nasal vowels. In *INTERSPEECH 2019* (pp. 3332–3336). Graz, Austria. doi: 10.21437/Interspeech.2019-2232

D'Alessandro, D., & Fougeron, C. (2018). Réduction de la coarticulation et vieillissement. In *Xxxiie journées d'études sur la parole* (pp. 410–418). Aix-en-Provence, France: ISCA. doi: 10.21437/JEP.2018-47

Das, B., Mandal, S., Mitra, P., & Basu, A. (2013). Effect of aging on speech features and phoneme recognition: A study on Bengali voicing vowels. *Int J Speech Technol*, *16*, 19–31. doi: 10.1007/s10772-012-9147-3

Davidson, L. (2006). Comparing tongue shapes from ultrasound imaging using smoothing spline analysis of variance. *J. Acoust. Soc. Am.*, *120*(1), 407–415. doi: 10.1121/1.2205133

Dawson, K. M. (2020). *Speech motor control in younger and older adults: The effect of age and individual differences on speech error production* (Doctoral dissertation). The City University of New York.

De Decker, P., & Mackenzie, S. (2017). Tracking the phonological status of /l/ in Newfoundland English: Experiments in articulation and acoustics. *J. Acoust. Soc. Am.*, *142*(1), 350–362. doi: 10.1121/1.4991349

De Decker, P., & Nycz, J. R. (2012). Are tense [æ]s really tense? The mapping between articulation and acoustics. *Lingua*, *122*(7), 810–821.

Decoster, W. (2000). Longitudinal voice changes: Facts and interpretation. *Journal of Voice*, *14*(2), 184–193.

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

Dehqan, A., Scherer, R. C., Dashti, G., Ansari-Moghaddam, A., & Fanaie, S. (2013). The effects of aging on acoustic parameters of voice. *Folia Phoniatrica et Logopaedica*, *64*(6), 265–270.

de Jong, N. H., & Wempe, T. (2009). Praat script to detect syllable nuclei and measure speech rate automatically. *Behavior Research Methods*, *41*(2), 385–390. doi: 10.3758/BRM.41.2.385

de Pinto, O., & Hollien, H. (1982). Speaking fundamental frequency characteristics of Australian women: Then and now. *Journal of Phonetics*, *10*(4), 367–375. doi: 10.1016/S0095-4470(19)31002-2

Desjardins, M., Halstead, L., Simpson, A., Flume, P., & Bonilha, H. S. (2022). The impact of respiratory function on voice in patients with presbyphonia. *Journal of Voice*, *36*(2), 256–271. doi: 10.1016/J.JVOICE.2020.05.027

Dey, S., & Alam, M. A. (2018). Formant based Bangla vowel perceptual space classification using Support Vector Machine and K-Nearest Neighbor method. In *21st International Conference of Computer and Information Technology (ICCIT)* (pp. 1–5). IEEE. doi: 10.1109/ICCITECHN.2018.8631948

D'haeseleer, E., Depypere, H., Claeys, S., Wuyts, F. L., Baudonck, N., & Van Lierde, K. M. (2011). Vocal characteristics of middle-aged premenopausal women. *Journal of Voice*, *25*(3), 360–366.

Diamond, G. M., Rochman, D., & Amir, O. (2010). Arousing primary vulnerable emotions in the context of unresolved anger: "Speaking about" versus "speaking to". *Journal of Counseling Psychology*, *57*(4), 402–410. doi: 10.1037/a0021115

Dietrich, M., Abbott, K. V., Gartner-Schmidt, J., & Rosen, C. A. (2008). The frequency of perceived stress, anxiety, and depression in patients with common pathologies affecting voice. *Journal of Voice*, *22*(4), 472–488.

Dimitrova, S., Andreeva, B., Gabriel, C., & Grünke, J. (2018). Speaker age effects on prosodic patterns in Bulgarian. In *Speech Prosody* (pp. 709–713). ISCA. doi: 10.21437/SpeechProsody.2018-144

Divenyi, P., Greenberg, S., & Meyer, G. (2006). *Dynamics of speech production and perception* (NATO Science Series, Ed.). IOS Press.

Dodig, S., Čepelak, I., & Pavić, I. (2019). Hallmarks of senescence and aging. *Biochemia medica*, *29*(3), 483–497. doi: 10.11613/BM.2019.030501

Doherty, T. J., Vandervoort, A. A., & Brown, W. F. (1993). Effects of ageing on the motor unit: A brief review. *Canadian Journal of Applied Physiology*, *18*(4), 331–358. doi: 10.1139/H93-029

Dokovova, M., Sabev, M., Scobbie, J. M., Lickley, R., & Cowen, S. (2019). Bulgarian vowel reduction in unstressed position: An ultrasound and acoustic investigation. In *19th ICPhS* (pp. 2720–2724).

Draxler, C., & Jänsch, K. (2004). SpeechRecorder - A universal platform independent multi-channel audio recording software. In *LREC'04* (pp. 559–562). Lisbon, Portugal.

Draxler, C., & Jänsch, K. (2017). *SpeechRecorder (3.12.0).*

Drioli, C., Tisato, G., Cosi, P., & Tesser, F. (2003). Emotions and voice quality: Experiments with sinusoidal modeling. In *ISCA Tutorial and Research Workshop on Voice Quality: Functions, Analysis and Synthesis.*

Dunashova, A. (2021). Intraspeaker variability of a professional lecturer: Ageing, genre, pragmatics vs. voice acting (Case Study). In *International Conference on Speech and Computer* (pp. 179–189).

Springer. doi: 10.1007/978-3-030-87802-3_17

Dyussenbayev, A. (2017). Age periods of Human Life. *Advances in Social Sciences Research Journal*, *4*(6). doi: 10.14738/assrj.46.2924

Eadie, T. L. (2000). Characteristics of the aging female voice caracteristiques de la voix feminine vieillissante. *La Revue D'Orthophonie et D'Audiologie*, *24*(4), 162–179.

Eichhorn, J. T., Kent, R. D., Austin, D., & 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. doi: 10.1016/J.JVOICE.2017.08.003

Ellgring, H., & Scherer, K. R. (1996). Vocal indicators of mood change in depression. *Journal of Nonverbal Behavior*, *20*(2), 83–110. doi: 10.1007/BF02253071

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

Elvira-García, W. (2014). *Visor-corrector de TextGrids. Script para Praat.* Barcelona: Laboratori de Fonètica. Retrieved from `https://github.com/wendyelviragarcia/TextGrid_annotation_and_modification_for_all_files_in_folder`

Endres, W., Bambach, W., & Flösser, G. (1971). Voice spectrograms as a function of age, voice disguise, and voice imitation. *J. Acoust. Soc. of Am.*, *49*(6B), 1842–1848.

Escudero, P., Boersma, P., Rauber, A. S., & Bion, R. A. H. (2009). A cross-dialect acoustic description of vowels: Brazilian and European Portuguese. *J. Acoust. Soc. Am.*, *126*(3), 1379–1393. doi: 10.1121/1.3180321

Esposito, A., Esposito, A. M., Likforman-Sulem, L., Maldonato, M. N., & Vinciarelli, A. (2016). On the significance of speech pauses in depressive disorders: results on read and spontaneous narratives. In A. Esposito et al. (Eds.), *Recent advances in nonlinear speech processing* (Vol. 48 ed., pp. 73–82). Springer. doi: 10.1007/978-3-319-28109-4_8

Ewald, O., Liina Asu, E., & Schötz, S. (2017). The formant dynamics of long close vowels in three varieties of Swedish. In *INTERSPEECH* (pp. 1412–1416). Stockholm, Sweden: ISCA. doi: 10.21437/Interspeech.2017-1134

Fant, G. (1970). *Acoustic theory of speech production: With calculations based on X-Ray studies of Russian articulations* (2nd ed.). Walter de Gruyter.

Feng, X., Todd, T., Hu, Y., Lintzenich, C. R., Carr, J. J., Browne, J. D., ... Butler, S. G. (2014). Age-related changes of hyoid bone position in healthy older adults with aspiration. *Laryngoscope*, *124*(6), E231–E236. doi: 10.1002/LARY.24453

Ferrand, C. T. (2002). Harmonics-to-Noise Ratio: An index of vocal aging. *Journal of Voice*, *16*(4), 480–487.

Ferreira, M. B., Carrilho, E., Lobo, M., Saramago, J., & Cruz, L. S. (1996). Variação linguística - Perspectiva dialectológica (Linguistic variation - dialectological perspective). In I. H. Faria, E. R. Pedro, I. Duarte, & C. A. M. Gouveia (Eds.), *Introdução à linguística geral e portuguesa* (pp. 479–502). Caminho.

Fikkert, P. (2005). From phonetic categories to phonological features specification: Acquiring the European Portuguese vowel system. *Lingue e linguaggio*, *4*(2), 263–280.

Fletcher, A. R., McAuliffe, M. J., Lansford, K. L., & 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.

Flint, A. J., Black, S. E., Campbell-Taylor, I., Gailey, G. F., & Levinton, C. (1993). Abnormal speech articulation, psychomotor retardation, and subcortical dysfunction in major depression. *Journal of Psychiatric Research*, *27*(3), 309–319. doi: 10.1016/0022-3956(93)90041-Y

Fortin, M., & Salgueiro, N. (1999). *O processo de investigação: da concepção à realização (The research process: from conception to realization)*. Lusociência - Edições Técnicas e Científicas, Lda.

Fougeron, C., & Audibert, N. (2011). Testing various metrics for the description of vowel distortion in dysarthria. In *ICPhS* (pp. 687–690).

Fougeron, C., Guitard-Ivent, F., & Delvaux, V. (2021). Multi-dimensional variation in adult speech as a function of age. *Languages*, *6*(4), 176. doi: 10.3390/LANGUAGES6040176

Fouquet, M., Pisanski, K., Mathevon, N., & Reby, D. (2016). Seven and up: Individual differences in male voice fundamental frequency emerge before puberty and remain stable throughout adulthood. *R. Soc. open sci.*, *3*, 1–9. doi: 10.1098/rsos.160395

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

Fox, R. A., & Jacewicz, E. (2010). Dialect and generational differences in vowel space areas. In *ExLing2010* (pp. 45–48). Athens, Greece: ISCA.

France, D. J., Shiavi, R. G., Silverman, S., Silverman, M., & Wilkes, M. (2000). Acoustical properties of speech as indicators of depression and suicidal risk. *IEEE Transactions on Biomedical Engineering*, *47*(7), 829–837. doi: 10.1109/10.846676

Fuchs, S., Gerstenberg, A., & Koenig, L. L. (2020). Changes in phonetic detail as a matter of discourse and aging: Evidence from a longitudinal study on French. In *SPPL2020: 2nd workshop on speech perception and production across the lifespan* (p. 6).

Fuchs, S., Koenig, L. L., & Gerstenberg, A. (2021). A longitudinal study of speech acoustics in older French females: Analysis of the filler particle euh across utterance positions. *Languages*, *6*(4), 211. doi: 10.3390/LANGUAGES6040211

Fuller, B. F., Horii, Y., & Conner, D. A. (1992). Validity and reliability of nonverbal voice measures as indicators of stressor-provoked anxiety. *Research in Nursing & Health*, *15*(5), 379–389. doi: 10.1002/nur.4770150507

Gahl, S., & Baayen, R. H. (2019). Twenty-eight years of vowels: Tracking phonetic variation through young to middle age adulthood. *Journal of Phonetics*, *74*, 42–54.

Galluzzi, F., & Garavello, W. (2018). The aging voice: A systematic review of presbyphonia. *European Geriatric Medicine*, *9*, 559–570. doi: 10.1007/s41999-018-0095-6

Georgeton, L., Antolík, T. K., & Fougeron, C. (2016). Effect of domain initial strengthening on vowel height and backness contrasts in French: Acoustic and ultrasound data. *J Speech Lang Hear Res*, *59*(6), S1575–S1586. doi: 10.1044/2016_JSLHR-S-15-0044

Gerstenberg, A., Fuchs, S., Kairet, J. M., Frankenberg, C., & Schröder, J. (2018). A cross-linguistic, longitudinal case study of pauses and interpausal units in spontaneous speech corpora of older speakers of German and French. In *Speech Prosody*. doi: 10.21437/SpeechProsody.2018-43

Gillam, R., Logan, K., & Pearson, N. (2009). *Test of childhood stuttering*. Austin, Texas: Pro-Ed Inc.

Girgus, J. S., & Yang, K. (2015). Gender and depression. *Current Opinion in Psychology*, *4*, 53–60. doi: 10.1016/j.copsyc.2015.01.019

Girgus, J. S., Yang, K., & Ferri, C. (2017). The gender difference in depression: Are elderly women at greater risk for depression than elderly men? *Geriatrics*, *2*(35), 1–21.

Goberman, A. M., Hughes, S., & Haydock, T. (2011). Acoustic characteristics of public speaking: Anxiety and practice effects. *Speech Communication*, *53*(6), 10–867.

Godfrey, H., & Knight, R. G. (1984). The validity of actometer and speech activity measures in the assessment of depressed patients. *The British Journal of Psychiatry*, *145*(2), 159–163.

Goldsmith, T. C. (2012). On the programmed/non-programmed aging controversy. *Biochemistry (Moscow)*, *77*(7), 729–732. doi: 10.1134/S000629791207005X

Gonçalves, T. M., Dos Santos, D. C., Pessin, A. B. B., & Martins, R. H. G. (2016, jun). Scanning electron microscopy of the presbylarynx. *Otolaryngology - Head and Neck Surgery*, *154*(6), 1073–1078. doi: 10.1177/0194599816637081

Gonzalez, S. (2021). Gridlines approach for dynamic analysis in speech ultrasound data: A multimodal app. *Laboratory Phonology: Journal of the Association for Laboratory Phonology*, *12*(1), 1–28. doi: 10.16995/labphon.6463

Goodglass, H., & Kaplan, E. (1983). *The assessment of aphasia and related disorders* (2nd ed.). Philadelphia, PA.: Lea and Febiger.

Goozée, J., Stephenson, D., Murdoch, B., Darnell, R., & Lapointe, L. (2005). Lingual kinematic strategies used to increase speech rate: Comparison between younger and older adults. *Clinical Linguistics & Phonetics*, *19*(4), 319–334.

Gorham-Rowan, M. M., & Laures-Gore, J. (2006). Acoustic-perceptual correlates of voice quality in elderly men and women. *Journal of Communication Disorders*, *39*(3), 171–184.

Goy, H., Fernandes, D. N., Pichora-Fuller, M. K., & van Lieshout, P. (2013). Normative voice data for younger and older adults. *Journal of Voice*, *27*(5), 545–555.

Goy, H., Kathleen Pichora-Fuller, M., & van Lieshout, P. (2016). Effects of age on speech and voice quality ratings. *J. Acoust. Soc. Am.*, *139*(4), 1648–1659. doi: 10.1121/1.4945094

Gracco, C., & Kahane, J. C. (1989). Age-related changes in the vestibular folds of the human larynx: A histomorphometric study. *Journal of Voice*, *3*(3), 204–212. doi: 10.1016/S0892-1997(89)80002-5

Greden, J. F. (1993). Psychomotor monitoring: A promise being fulfilled? *Journal of Psychiatric Research*, *27*(3), 285–287. doi: 10.1016/0022-3956(93)90039-5

Greden, J. F., Albala, A. A., Smokler, I. A., Gardner, R., & Carroll, B. J. (1981). Speech pause time: a marker of psychomotor retardation among endogenous depressives. *Biological Psychiatry*, *16*(9), 851–859.

Guimarães, I. (2007). *A ciência e a arte da voz humana (The science and art of the human voice)* (1st ed.). Alcabideche: Escola Superior de Saúde de Alcoitão.

Guimarães, I., & Abberton, E. (2005). Fundamental frequency in speakers of Portuguese for different voice samples. *Journal of Voice*, *19*(4), 592–606.

Hagen, P., Lyons, G. D., & Nuss, D. W. (1996). Dysphonia in the elderly: Diagnosis and management of age-related voice changes. *Southern Medical Journal*, *89*(2), 204–207.

Hagenaars, M. A., & van Minnen, A. (2005). The effect of fear on paralinguistic aspects of speech in patients with panic disorder with agoraphobia. *Journal of Anxiety Disorders*, *19*(5), 521–537. doi: 10.1016/J.JANXDIS.2004.04.008

Hall, J. A., Harrigan, J. A., & Rosenthal, R. (1995). Nonverbal behavior in clinician - patient interaction. *Applied & Preventive Psychology*, *4*, 21–37. doi: 10.1016/S0962-1849(05)80049-6

Hämäläinen, A., Avelar, J., Rodrigues, S., Dias, M. S., Kolesinski, A., Fegyó, T., ... Hewson, D. (2014). The EASR corpora of European Portuguese, French, Hungarian and Polish elderly speech. In *LREC* (pp. 1458–1464).

Hämäläinen, A., Pinto, F. M., Dias, M. S., Júdice, A., Pires, C. G., Teixeira, V. D., ... 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.

Hammond, T. H., Gray, S. D., & Butler, J. E. (2000). Age- and gender-related collagen distribution in human vocal folds. *Ann Otol Rhinol Laryngol*, *109*, 913–920. doi: 10.1177/000348940010901004

Hardcastle, W. J., Laver, J., & Gibbon, F. E. (2012). *The handbook of phonetic sciences* (Vol. 116). John Wiley & Sons.

Hardy, P., Jouvent, R., & Widlöcher, D. (1984). Speech pause time and the retardation rating scale for depression (ERD): Towards a reciprocal validation. *Journal of Affective Disorders*, *6*(1), 123–127. doi: 10.1016/0165-0327(84)90014-4

Harrington, J., Cassidy, S., & Cassidy, S. (1999). *Techniques in speech acoustics* (Vol. 8). Springer Science & Business Media.

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

Hartman, D. E., & Danhauer, J. L. (1976). Perceptual features of speech for males in four perceived age decades. *J. Acoust. Soc. Am*, *59*(3), 713–715.

Hazan, V. (2017). Speech communication across the lifespan. *Acoustics Today*, *13*(1), 36–43.

Hazan, V., Tuomainen, O., Kim, J., Davis, C., Sheffield, B., & Brungart, D. (2018). Clear speech adaptations in spontaneous speech produced by young and older adults. *J. Acoust. Soc. Am.*, *144*(3), 1331–1346.

He, W., Goodkind, D., & Kowal, P. R. (2016). *An aging world: 2015* (Vol. P95/16-1; Tech. Rep.). Washington, DC.

Hermes, A., Bourbon, A., & Cecile, F. (2020). Aging effects on prosodic structuring in French. In *SPPL2020: 2nd workshop on speech perception and production across the lifespan* (p. 119).

Hermes, A., Mertens, J., & Mücke, D. (2018). Age-related effects on sensorimotor control of speech production. In *INTERSPEECH* (pp. 1526–1530). Hyderabad.

Higgins, M. B., & Saxman, J. H. (1991). A comparison of selected phonatory behaviors of healthy aged and young adults. *J Speech Hear Res*, *34*(5), 1000–1010.

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

Hirano, M., Kurita, S., & Sakaguchi, S. (1989). Ageing of the vibratory tissue of human vocal folds. *Acta Otolaryngol (Stockh)*, *107*, 428–433. doi: 10.3109/00016488909127535

Hixon, T., Weismer, G., & Hoit, J. (2018). *Preclinical speech science: Anatomy, physiology, acoustics, and perception* (3rd ed.). Plural Publishing.

Hoffnung, M., Hoffnung, R. J., Seifert, K. L., Hine, A., Pausé, C., Ward, L., … Smith, R. B. (2015). *Lifespan development: A chronological approach* (4th ed.). John Wiley & Sons Australia, Ltd.

Hoit, J. D., & Hixon, T. J. (1987). Age and speech breathing. *J Speech Hear Res*, *30*(3), 351–366.

Hoit, J. D., Hixon, T. J., Altman, M. E., & Morgan, W. J. (1989). Speech breathing in women. *J Speech Hear Res*, *32*(2), 353–365.

Hoit, J. D., Watson, P. J., Hixon, K. E., McMahon, P., & Johnson, C. L. (1994). Age and velopharyngeal function during speech production. *J Speech Hear Res*, *37*(2), 295–302.

Hollien, H., & Shipp, T. (1972). Speaking fundamental frequency and chronologic age in males. *J Speech Hear Res*, *15*(1), 155–159.

Holloszy, J. O. (2000). The biology of aging. *Mayo Clinic Proceedings*, *75*(1), S3–S9. doi: 10.1016/S0025-6196(19)30634-2

Hönig, F., Batliner, A., Nöth, E., Schnieder, S., & Krajewski, J. (2014). Automatic modelling of depressed speech: Relevant features and relevance of gender. In *INTERSPEECH* (pp. 1248–1252). Singapore: ISCA.

Hooper, C. R., & Cralidis, A. (2009). Normal changes in the speech of older adults: You've still got what it takes; it just takes a little longer! *Perspectives on Gerontology*, *14*(2), 47–56. doi: 10.1044/gero14.2.47

Hopkins, W. G. (2002). *A scale of magnitudes for effect statistics. A new view of statistics.* Retrieved from `www.sportsci.org/resource/stats/effectmag.html`

Horwitz, R., Quatieri, T. F., Helfer, B. S., Yu, B., Williamson, J. R., & Mundt, J. (2013). On the relative importance of vocal source, system, and prosody in human depression. In *2013 IEEE International Conference on Body Sensor Networks* (pp. 1–6). IEEE.

Huber, J. E. (2008). Effects of utterance length and vocal loudness on speech breathing in older adults. *Respiratory Physiology \&; Neurobiology*, *164*(3), 323–330. doi: 10.1016/j.resp.2008.08.007

Huber, J. E., & Spruill III, J. (2008). Age-related changes to speech breathing with increased vocal loudness. *J Speech Lang Hear Res*, *51*(3), 651–668. doi: 10.1044/1092-4388(2008/047)

Hutchinson, J. M., Robinson, K. L., & Nerbonne, M. A. (1978). Patterns of nasalance in a sample of normal gerontologic subjects. *Journal of Communication Disorders*, *11*(6), 469–481. doi: 10.1016/0021-9924(78)90021-7

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

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

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

Jacewicz, E., Fox, R. A., & 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.

Jacewicz, E., Fox, R. A., & Wei, L. (2010). Between-speaker and within-speaker variation in speech tempo of American English. *J. Acoust. Soc. Am.*, *128*(2), 839–850.

Jannedy, S., & Weirich, M. (2017). Spectral moments vs discrete cosine transformation coefficients: Evaluation of acoustic measures distinguishing two merging German fricatives. *J. Acoust. Soc. Am.*, *142*(1), 395–405. doi: 10.1121/1.4991347

Janssens, J.-P. (2005). Aging of the respiratory system: impact on pulmonary function tests and adaptation to exertion. *Clinics in chest medicine*, *26*(3), 469–484.

Janssens, J. P., Pache, J. C., & Nicod, L. P. (1999). Physiological changes in respiratory function associated with ageing. *European Respiratory Journal*, *13*(1), 197–205.

Jayakumar, T., Benoy, J. J., & Yasin, H. M. (2020). Effect of age and gender on acoustic Voice Quality Index across lifespan: A cross-sectional study in Indian population. *Journal of Voice*, in press. doi: 10.1016/j.jvoice.2020.05.025

Jin, K. (2010). Modern biological theories of aging. *Aging and Disease*, *1*(2), 72–74. doi: 10.1093/jn/119.6.952

Jin, S.-H., & 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. doi: 10.1121/1.4798620

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

Jorm, A. F. (2000). Does old age reduce the risk of anxiety and depression? A review of epidemiological studies across the adult life span. *Psychological Medicine*, *30*, 11–22.

Kafková, M. P. (2016). The ”real” old age and the transition between the third and fourth age. *Sociológia*, *48*(6), 622–640.

Kalmovich, L. M., Elad, D., Zaretsky, U., Adunsky, A., Chetrit, A., Sadetzki, S., … Wolf, M. (2005). Endonasal geometry changes in elderly people: Acoustic rhinometry measurements. *Journal of Gerontology: MEDICAL SCIENCES*, *60A*(3), 396–398.

Kaur, J., & Narang, V. (2015). Variation of pitch and formants in different age group. *Int J Multidiscip Res Mod Educ*, *1*, 517–521.

Kendall, K. (2007). Presbyphonia: a review. *Current opinion in otolaryngology & head and neck surgery*, *15*(3), 137–140.

Kent, R. D., & Read, C. (2002). *The acoustic analysis of speech* (2a ed.). Australia; Canada; Mexico: Singular, Thomson Learning.

Kent, R. D., & Rountrey, C. (2020). What acoustic studies tell us about vowels in developing and disordered speech. *AJSLP*, *29*, 1749–1778.

Kent, R. D., & Vorperian, H. K. (2018). Static measurements of vowel formant frequencies and bandwidths: A review. *Journal of Communication Disorders*, *74*, 74–97.

Kersing, W., & Jennekens, F. G. (2004). Age-related changes in human thyroarytenoid muscles: a histological and histochemical study. *Eur Arch Otorhinolaryngol*, *261*(7), 386–392. Retrieved from `http://www.ncbi.nlm.nih.gov/pubmed/14605813` doi: 10.1007/s00405-003-0702-z

Kessler, R. C., Birnbaum, H. G., Shahly, V., Bromet, E., Hwang, I., McLaughlin, K. A., ... Stein, D. J. (2010). Age differences in the prevalence and co-morbidity of DSM-IV major depressive episodes: results from the WHO World Mental Health Survey Initiative. *Depression and Anxiety*, *27*(4), 351–364. doi: 10.1002/da.20634

Kirkham, S., & Nance, C. (2017). An acoustic-articulatory study of bilingual vowel production: Advanced tongue root vowels in Twi and tense/lax vowels in Ghanaian English. *Journal of Phonetics*, *62*, 65–81.

Kirkwood, T. B., & Melov, S. (2011). On the programmed/non-programmed nature of ageing within the life history. *Current Biology*, *21*(18), 701–707. doi: 10.1016/J.CUB.2011.07.020

Kisler, T., Reichel, U., & Schiel, F. (2017). Multilingual processing of speech via web services. *Computer Speech and Language*, *45*, 326–347. doi: 10.1016/j.csl.2017.01.005

Klein, D. R. (1980). Oral soft tissue changes in geriatric patients. *Bull. N.Y. Acad. Med.*, *56*(8), 721–727.

Klumpp, H., & Deldin, P. (2010). Review of brain functioning in depression for semantic processing and verbal fluency. *International Journal of Psychophysiology*, *75*(2), 77–85. doi: 10.1016/j.ijpsycho.2009.10.003

Km, J.-E., & Yoon, K. (2012). An analysis of the vowel formants of the young versus old speakers in the Buckeye corpus. *Phonetics and Speech Sciences*, *4*(4), 29–35.

Kost, K. M., & Sataloff, R. T. (2020). The aging voice. *Current Otorhinolaryngology Reports*, *8*(3), 321–328. doi: 10.1007/S40136-020-00285-8

Kraepelin, E. (1921). Manic depressive insanity and paranoia. *The Journal of Nervous and Mental Disease*, *53*(4).

Kuehner, C. (2017). Why is depression more common among women than among men? *The Lancet Psychiatry*, *4*(2), 146–158. doi: 10.1016/S2215-0366(16)30263-2

Kuny, S., & Stassen, H. (1993). Speaking behavior and voice sound characteristics in depressive patients during recovery. *Journal of Psychiatric Research*, *27*(3), 289–307. doi: 10.1016/0022-3956(93)90040-9

Kuruvilla-Dugdale, M., Dietrich, M., McKinley, J. D., & Deroche, C. (2020). An exploratory model of speech intelligibility for healthy aging based on phonatory and articulatory measures. *Journal of Communication Disorders*, *87*, 105995. doi: 10.1016/J.JCOMDIS.2020.105995

Kyriaki, D. (2021). *Speech production across ages: An acoustic analysis* (Master's thesis). Aristotle University of Thessaloniki.

Lã, F. M., & Ardura, D. (2020). What voice-related metrics change with menopause? A systematic review and meta-analysis study. *Journal of Voice*. doi: 10.1016/j.jvoice.2020.06.012

Ladefoged, P., & Johnson, K. (2011). *A course in Phonetics* (6th ed.). Michael Rosenberg.

Lalley, P. M. (2013). The aging respiratory system-Pulmonary structure, function and neural control. *Respiratory Physiology and Neurobiology*, *187*(3), 199–210. doi: 10.1016/j.resp.2013.03.012

Lancia, L., Rausch, P., & Morris, J. S. (2015). Automatic quantitative analysis of ultrasound tongue contours via wavelet-based functional mixed models. *J. Acoust. Soc. Am.*, *137*(2), EL178–EL183. doi: 10.1121/1.4905881

Lanitis, A. (2010). A survey of the effects of aging on biometric identity verification. *International*

Journal of Biometrics, *2*(1), 34–52.

Laukka, P., Linnman, C., Åhs, F., Pissiota, A., Frans, Ö., Faria, V., ... Furmark, T. (2008). In a nervous voice: Acoustic analysis and perception of anxiety in social phobics' speech. *Journal of Nonverbal Behavior*, *32*(4), 195.

Lee, S.-H., Yu, J.-F., Hsieh, Y.-H., & Lee, G.-S. (2015). Relationships between formant frequencies of sustained vowels and tongue contours measured by ultrasonography. *Am J Speech Lang Pathol*, *24*(4), 739–749.

Lee, W.-S. (2016). Articulatory–acoustical relationship in Cantonese vowels. *Language and Linguistics*, *17*(4), 477–500. doi: 10.1177/1606822X16637058

Lemoine, M. (2020). Defining aging. *Biology and Philosophy*, *35*, 46. doi: 10.1007/S10539-020-09765-Z

Lenell, C., Sandage, M. J., & Johnson, A. M. (2019). A tutorial of the effects of sex hormones on laryngeal senescence and neuromuscular response to exercise. *J Speech Lang Hear Res*, *62*, 602–6102. doi: 10.1044/2018_JSLHR-S-18-0179

Leung, Y., Oates, J., Papp, V., & Chan, S.-P. (2020). Formant frequencies of adult speakers of Australian English and effects of sex, age, geographical location, and vowel quality. *Journal of Voice*, (in press). doi: 10.1016/j.jvoice.2020.09.026

Leung, Y., Oates, J., Papp, V., & Chan, S. P. (2022). Speaking fundamental frequencies of adult speakers of Australian English and effects of sex, age, and geographical location. *Journal of Voice*, *36*(3), 434.e1–434.e15. doi: 10.1016/J.JVOICE.2020.06.014

Lieberman, P., & Biumstein, S. E. (1988). *Speech physiology, speech perception, and acoustic phonetics (Cambridge Studies in Speech Science and Communication)*. Cambridge University Press.

Linville, S. E. (1992). Glottal gap configurations in two age groups of women. *J Speech Hear Res*, *35*(6), 1209–1215.

Linville, S. E. (1996). The sound of senescence. *Journal of Voice*, *10*(2), 190–200.

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

Linville, S. E., & Fisher, H. B. (1985). Acoustic characteristics of women's voices with advancing age. *Journal of Gerontology*, *40*(3), 324–330. doi: 10.1093/geronj/40.3.324

Linville, S. E., & Rens, J. (2001). Vocal tract resonance analysis of aging voice using long-term average spectra. *Journal of Voice*, *15*(3), 323–330.

Liss, J. M., Weismer, G., & Rosenbek, J. C. (1990). Selected acoustic characteristics of speech production in very old males. *Journal of Gerontology*, *45*(2), P35–P45.

López-Otín, C., Blasco, M. A., Partridge, L., Serrano, M., & Kroemer, G. (2013). The hallmarks of aging. *Cell*, *153*(6), 1194–1217. doi: 10.1016/j.cell.2013.05.039

Lortie, C. L., Thibeault, M., Guitton, M. J., & Tremblay, P. (2015). Effects of age on the amplitude, frequency and perceived quality of voice. *Age*, *37*(6), 117.

Low, D. M., Bentley, K. H., & Ghosh, S. S. (2020). Automated assessment of psychiatric disorders using speech: A systematic review. *Laryngoscope Investigative Otolaryngology*, *5*(1), 96–116. doi: 10.1002/lio2.354

Lutzross, A., Schuerman, W., Sprouse, R., & Gahl, S. (2013). Development of vowel spaces from age

21 to age 49 in a group of 8 talkers. *J. Acoust. Soc. Am.*, *134*(5), 4107. doi: 10.1121/1.4831073

Ma, E., & Love, A. (2010). Electroglottographic evaluation of age and gender effects during sustained phonation and connected speech. *Journal of Voice*, *24*(2), 146–152. doi: 10.1016/j.jvoice.2008.08.004

Madureira, S. (2016). Intonation and variation: The multiciplicity of forms and senses. *Dialectologia. Special issue*, *VI*, 57–74.

Mahne, A., El-Haddad, G., Alavi, A., Houseni, M., Moonis, G., Mong, A., ... Torigian, D. A. (2007). Assessment of age-related morphological and functional changes of selected structures of the head and neck by computed tomography, magnetic resonance imaging, and positron emission tomography. *Seminars in Nuclear Medicine*, *37*(2), 88–102. doi: 10.1053/J.SEMNUCLMED.2006.10.003

Makiyama, K., & Hirano, S. (2017). *Aging voice.* Singapore: Springer.

Mallick, A., Garas, G., & McGlashan, J. (2019). Presbylaryngis: a state-of-the-art review. *Current Opinion in Otolaryngology and Head and Neck Surgery*, *27*(3), 168–177.

Marino, V. C., Cardoso, V. M., de Boer, G., Dutka, J. d., Fabbron, E. M., & Bressmann, T. (2018). Normative nasalance scores for middle-aged and elderly speakers of Brazilian Portuguese. *Folia Phoniatrica et Logopaedica*, *70*(2), 82–89. doi: 10.1159/000489334

Markó, A., & Bóna, J. (2010). Fundamental frequency patterns: The factors of age and speech type. In *Sociophonetics, at the crossroads of speech variation, processing and communication.* Pisa.

Martin, J. H., Diamond, B., Aviv, J. E., Jones, M. E., Keen, M. S., Wee, T. A., & Blitzer, A. (1994). Age-related changes in pharyngeal and supraglottic sensation. *Ann Otol Rhinol Laryngol*, *103*, 749–752. doi: 10.1177/000348949410301001

Martins, M. R. D. (1973). Análise acústica das vogais orais tónicas em Português (Acoustic analysis of stressed oral vowels in Portuguese). *Boletim de Filologia*, *22*, 303–314.

Martins, R., Benito Pessin, A., Nassib, D., Branco, A., Rodrigues, S., & Matheus, S. (2015). Aging voice and the laryngeal muscle atrophy. *The Laryngoscope*, *125*(11), 2518–2521. doi: 10.1002/lary.25398

Martins, V., & Andrade, C. (2011). Estudo das pausas em idosos (Study of pauses in elderly). *Revista da Sociedade Brasileira de Fonoaudiologia*, *16*(3), 344–349. doi: 10.1590/S1516-80342011000300017

Massimo, P., & Elisa, P. (2014). Age and Rhtymic Variations: A study on Italian. In *INTERSPEECH 2014* (pp. 1234–1237). Singapore: ISCA.

Mateus, M. H., & D'Andrade, E. (2000). *The phonology of Portuguese*. Oxford: Oxford University Press.

Matias, M. (1993). *Aspectos da estrutura sociolinguística da cidade de Aveiro (Aspects of the sociolinguistic structure of Aveiro city)* (Doctoral dissertation). Universidade de Coimbra.

Maurer, D. (2016). *Acoustics of the vowel: Preliminaries*. PETER LANG LTD International Academic Publishers.

Mautner, H. (2011). *A cross-system instrumental voice profile of the aging voice: With considerations of jaw posture effects* (Doctoral dissertation). University of Canterbury, New Zealand.

McCloy, D., Wright, R., & Souza, P. (2014). Modeling intrinsic intelligibility variation: vowel-space size and structure. In *Meetings on acoustics* (Vol. 18, p. 60007). ASA.

McDougall, K., & Nolan, F. (2007). Discrimination of speakers using the formant dynamics of /u:/ in British English. In *ICPhS XVI* (pp. 1825–1828). Saarbrücken, Germany.

Ménard, L., Toupin, C., Baum, S. R., Drouin, S., Aubin, J., & Tiede, M. (2013). Acoustic and articulatory analysis of French vowels produced by congenitally blind adults and sighted adults. *J. Acoust. Soc. Am.*, *134*(4), 2975–2987.

Mendes, A., Guerreiro, D., Simões, M., & Moreira, M. (2013). *Fisiologia da Técnica Vocal (Physiology of the vocal technique)*. Lisboa: Lusociência.

Mertens, J., Mücke, D., & 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.

Michalik, M., Kaczorowska-Bray, K., Milewski, S., & Solak, A. (2018). The speech rate of people at an advanced old age. *Acta Neuropsychologica*, *16*(2), 167–175. doi: 10.5604/01.3001.0012.1970

Mielke, J. (2015). An ultrasound study of Canadian French rhotic vowels with polar smoothing spline comparisons. *J. Acoust. Soc. Am.*, *137*(5), 2858–2869. doi: 10.1121/1.4919346

Mielke, J., Carignan, C., & Thomas, E. R. (2017). The articulatory dynamics of pre-velar and pre-nasal/æ/-raising in English: An ultrasound study. *J. Acoust. Soc. Am.*, *142*(1), 332–349.

Mifune, E., Justino, V. S. S., Camargo, Z., & Gregio, F. (2007). Análise acústica da voz do idoso: Caracterização da freqüência fundamental (Acoustic analysis of aging voice: fundamental frequency characterization). *Revista CEFAC*, *9*(2), 238–247.

Mizuta, M. (2017, jul). Age-related histological changes of the vocal folds. In K. Makiyama & S. Hirano (Eds.), *Aging voice* (pp. 9–25). Singapore: Springer.

Morgan, E. E., & Rastatter, M. (1986). Variability of voice fundamental frequency in elderly female speakers. *Perceptual and motor skills*, *63*(1), 215–218.

Morris, R. J. (1986). *Age-related differences in articulatory physiology among adult females* (Doctoral dissertation). University of Florida.

Morris, R. J., & Brown, W. S. (1994). Age-related differences in speech variability among women. *Journal of Communication Disorders*, *27*(1), 49–64. doi: 10.1016/0021-9924(94)90010-8

Morris, R. J., Brown Jr, W. S., Hicks, D. M., & Howell, E. (1995). Phonational profiles of male trained singers and nonsingers. *Journal of Voice*, *9*(2), 142–148.

Morrison, G. S. (2013). Vowel inherent spectral change in forensic voice comparison. In G. S. Morrison & P. F. Assmann (Eds.), *Vowel inherent spectral change* (pp. 263–282). Springer Berlin Heidelberg. doi: 10.1007/978-3-642-14209-3_11

Morrison, G. S., & Assmann, P. F. (2013). Vowel inherent spectral change. In *Modern acoustics and signal processing.* Berlin: Springer.

Mortensen, L., Meyer, A. S., & Humphreys, G. W. (2006). Age-related effects on speech production: A review. *Lang Cognitive Proc*, *21*(1-3), 238–290.

Mortimore, I. L., Bennett, S. P., & Douglas, N. J. (2000). Tongue protrusion strength and fatiguability: Relationship to apnoea/hypopnoea index and age. *J. Sleep Res.*, *9*(4), 389–393. doi: 10.1046/J.1365-2869.2000.00222.X

Mortimore, I. L., Fiddes, P., Stephens, S., & Douglas, N. J. (1999). Tongue protrusion force and fatiguability in male and female subjects. *European Respiratory Journal*, *14*, 191–195.

Mou, Z., Teng, W., Ouyang, H., Chen, Y., Liu, Y., Jiang, C., ... Chen, Z. (2019). Quantitative analysis of vowel production in cerebral palsy children with dysarthria. *Journal of Clinical Neuroscience*, *66*, 77–82. doi: 10.1016/j.jocn.2019.05.020

Mozaffari, M. H., & Lee, W.-S. (2019). Domain adaptation for ultrasound tongue contour extraction using transfer learning: A deep learning approach. *J. Acoust. Soc. Am.*, *146*(5), EL431–EL437. doi: 10.1121/1.5133665

Mozaffari, M. H., Wen, S., Wang, N., & Lee, W. (2019). Real-time automatic tongue contour tracking in ultrasound video for guided pronunciation training. In *14th International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications (VISIGRAPP 2019)* (Vol. 1, pp. 302–309). doi: 10.5220/0007523503020309

Mücke, D., Thies, T., Mertens, J., & Hermes, A. (2020). Age-related effects of prosodic prominence in vowel articulation. In *12th International Seminar on Speech Production (ISSP).*

Mueller, P. B. (1997). The aging voice. *Seminars in Speech and Language*, *18*(2), 159–168. doi: 10.1055/S-2008-1064070

Mundt, J. C., Snyder, P. J., Cannizzaro, M. S., Chappie, K., & Geralts, D. S. (2007). Voice acoustic measures of depression severity and treatment response collected via interactive voice response (IVR) technology. *J of neurolinguistics*, *20*(1), 50–64. doi: 10.1016/j.jneuroling.2006.04.001

Mundt, J. C., Vogel, A. P., Feltner, D. E., & Lenderking, W. R. (2012). Vocal acoustic biomarkers of depression severity and treatment response. *Biological Psychiatry*, *72*(7), 580–587.

Munson, B. (2007). Lexical access, lexical representation, and vowel production. *Laboratory phonology*, *9*, 201–228.

Munson, B., & Solomon, N. P. (2004). The effect of phonological neighborhood density on vowel articulation. *J Speech Lang Hear Res*, *47*(5), 1048–1058.

Murray, I. R., & Arnott, J. L. (1993). Toward the simulation of emotion in synthetic speech: A review of the literature on human vocal emotion. *J. Acoust. Soc. Am.*, *93*(2), 1097–1108. doi: 10.1121/1.405558

Murry, T., Brown, W. S., & Morris, R. J. (1995). Patterns of fundamental frequency for three types of voice samples. *Journal of Voice*, *9*(3), 282–289.

Mwangi, S., Spiegl, W., Hoenig, F., Haderlein, T., Maier, A., & Noeth, E. (2009). Effects of vocal aging on fundamental frequency and formants. In *International Conference on Acoustics NAG/DAGA* (pp. 1761–1764). Rotterdam.

Nearey, T. M., & Assmann, P. F. (1986). Modeling the role of inherent spectral change in vowel identification. *J. Acoust. Soc. Am.*, *80*(5), 1297–1308. doi: 10.1121/1.394433

Neel, A. T. (2008). Vowel space characteristics and vowel identification accuracy. *J Speech Lang Hear Res*, *51*(3), 574–585.

Neel, A. T., & Palmer, P. M. (2012). Is tongue strength an important influence on rate of articulation in diadochokinetic and reading tasks? *J Speech Lang Hear Res*, *55*, 235–246. doi: 10.1044/1092-4388(2011/10-0258)

Neugarten, B. (1974). Age groups in American society and the rise of the young-old. *The ANNALS of the American Academy of Political and Social Science*, *415*(1), 187–198.

Nilsonne, Å. (1987). Acoustic analysis of speech variables during depression and after improvement. *Acta Psychiatrica Scandinavica*, *76*(3), 235–245.

Nishio, M., & Niimi, S. (2008). Changes in speaking fundamental frequency characteristics with aging. *Folia Phoniatrica et Logopaedica*, *60*(3), 120–127.

Noble, A., Boukerroui, D., Noble, A., Boukerroui, D., Noble, J. A., Member, S., & Boukerroui, D. (2006). Ultrasound image segmentation: A survey. *IEEE Transactions on Medical Imaging*, *25*, 987–1010.

Nooteboom, S. (1997). The prosody of speech: Melody and rhythm. In *The handbook of phonetic sciences* (pp. 640–673).

Nosraty, L. (2018). *Successful aging among the oldest old* (Doctoral dissertation). University of Tampere.

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

Oliveira, C., Martins, P., Silva, S., & Teixeira, A. (2012). An MRI study of the oral articulation of European Portuguese nasal vowels. In *13th INTERSPEECH* (pp. 2690–2693). Portland, OR, USA.

Oliveira, C., Martins, P., & Teixeira, A. (2009). Speech rate effects on European Portuguese nasal vowels. In *INTERSPEECH* (pp. 480–483). Brighton.

Ozdas, A., Shiavi, R. G., Silverman, S. E., Silverman, M. K., & Wilkes, D. M. (2004). Investigation of vocal jitter and glottal flow spectrum as possible cues for depression and near-term suicidal risk. *IEEE Transactions on Biomedical Engineering*, *51*(9), 1530–1540. doi: 10.1109/TBME.2004.827544

Özseven, T., Düğenci, M., Doruk, A., & Kahraman, H. (2018). Voice traces of anxiety: Acoustic parameters affected by anxiety disorder. *Archives of Acoustics*, *43*(4), 625–636.

Pais-Ribeiro, J., Silva, I., Ferreira, T., Martins, A., Meneses, R., & Baltar, M. (2007). Validation study of a Portuguese version of the Hospital Anxiety and Depression Scale. *Psychology, health & medicine*, *12*(2), 225–237.

Pakhomov, S. V., Kaiser, E. A., Boley, D. L., Marino, S. E., Knopman, D. S., & Birnbaum, A. K. (2011). Effects of age and dementia on temporal cycles in spontaneous speech fluency. *Journal of Neurolinguistics*, *24*(6), 619–635. doi: 10.1016/j.jneuroling.2011.06.002

Palao, S. (2017). *Color pictograms ARASAAC.* Aragonese Portal of Augmentative and Alternative Communication: Creative Commons' License (BY-NC-SA). Retrieved from `http://arasaac.org`

Papalia, D. E., Feldman, R. D., Monteiro, C., & Silva, M. d. C. (2013). *Desenvolvimento humano (Experience human development)* (12th ed.). Porto Alegre: AMGH Editora.

Paulsen, F. P., & Tillmann, B. N. (1998). Degenerative changes in the human cricoarytenoid joint. *Archives of Otolaryngology—Head & Neck Surgery*, *124*(8), 903.

Pegoraro-Krook, M. I. (1988). Speaking fundamental frequency characteristics of normal Swedish subjects obtained by glottal frequency analysis. *Folia Phoniatrica et Logopaedica*, *40*(2), 82–90. doi: 10.1159/000265888

Pellegrini, T., Hämäläinen, A., de Mareüil, P. B., Tjalve, M., Trancoso, I., Candeias, S., ... 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* (pp. 852–856). Lyon.

Pellegrini, T., Hedayati, V., Trancoso, I., Hämäläinen, A., & Dias, M. S. (2014). Speaker age estimation for elderly speech recognition in European Portuguese. In *INTERSPEECH* (pp. 2962–2966).

Pellegrino, E. (2019). The effect of healthy aging on within-speaker rhythmic variability: A case study on Noam Chomsky. *Loquens*, *6*(1), e060. doi: 10.3989/loquens.2019.060

Pellegrino, E., He, L., & Dellwo, V. (2018). The effect of ageing on speech rhythm: A study on Zurich German. In *Speech Prosody* (pp. 133–137).

Pellegrino, E., He, L., & Volker, D. (2021). Age-related rhythmic variations: The role of syllable intensity variability. *TRANEL - Travaux neuchâtelois de linguistique*, *74*, 167–185. doi: 10.5167/uzh-203680

Pérez, X. Á. (2014). European Portuguese dialectal features: A comparison with Cintra's proposal. *Journal of Portuguese Linguistics*, *13*(1), 29–62.

Pessin, A., Tavares, E., Gramuglia, A., de Carvalho, L., & Martins, R. (2017). Voice and ageing: Clinical, endoscopic and acoustic investigation. *Clinical Otolaryngology*, *42*(2), 330–335. doi: 10.1111/coa.12725

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

Pontes, P., Brasolotto, A., & Behlau, M. (2005). Glottic characteristics and voice complaint in the elderly. *Journal of Voice*, *19*(1), 84–94.

Pontes, P., Yamasaki, R., & Behlau, M. (2006). Morphological and functional aspects of the senile larynx. *Folia Phoniatrica et Logopaedica*, *58*, 151–158. doi: 10.1159/000091729

Prieto, P., & Roseano, P. (2018). Prosody: Stress, rhythm, and intonation. In K. Geeslin (Ed.), *The cambridge handbook of spanish linguistics* (pp. 211–236). Cambridge: Cambridge University Press. doi: 10.1017/9781316779194.011

Protopapas, A., & Lieberman, P. (1997). Fundamental frequency of phonation and perceived emotional stress. *J. Acoust. Soc. Am.*, *101*(4), 2267–2277. doi: 10.1121/1.418247

Ptacek, P. H., Sander, E. K., Maloney, W. H., & Jackson, C. C. R. (1966). Phonatory and related changes with advanced age. *J Speech Hear Res*, *9*(3), 353–360.

Quadagno, J. (2017). *Aging and the life course: An introduction to social gerontology* (7th ed.). McGraw-Hill Higher Education.

Quatieri, T. F., & Malyska, N. (2012). Vocal-source biomarkers for depression: a link to psychomotor activity. In *INTERSPEECH* (pp. 1059–1062).

Quené, H. (2013). Longitudinal trends in speech tempo: The case of Queen Beatrix. *J. Acoust. Soc. Am.*, *133*(6), EL452–EL457. doi: 10.1121/1.4802892

Radisic, M. (2014). *An ultrasound and acoustic study of Turkish rounded/unrounded vowel pairs* (Doctoral dissertation). University of Toronto.

Ramig, L., Gray, S., Baker, K., Corbin-Lewis, K., Buder, E., Luschei, E., ... Smith, M. (2001). The aging voice: A review, treatment data and familial and genetic perspectives. *Folia Phoniatrica* *et Logopaedica*, *53*(5), 252–265.

Ramig, L., & Ringel, R. (1983). Effects of physiological aging on selected acoustic characteristics of voice. *J Speech Hear Res*, *26*(1), 22–30.

Raphael, L. J., Borden, G. J., & Harris, K. S. (2011). *Speech science primer* (6th ed.). Lippincott Williams & Wilkins.

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

Rato, A., Rodrigues, C., & Varanda, P. (2017). Stressed vowels of European Portuguese in spontaneous speech. In P. Barbosa, M. d. C. de Paiva, & C. Rodrigues (Eds.), *Studies on variation in portuguese* (chap. 1). JOHN BENJAMINS PUBLISHING CO.

Raz, N., & Rodrigue, K. (2006). Differential aging of the brain: Patterns, cognitive correlates and modifiers. *Neuroscience and Biobehavioral Reviews*, *30*, 730–748. doi: 10.1016/j.neubiorev.2006.07.001

Redford, M. (2015). *The handbook of speech production*. Wiley Blackwell.

Reubold, U., & Harrington, J. (2015). Disassociating the effects of age from phonetic change. *Language Development: The life span perspective*, 9–37.

Reubold, U., Harrington, J., & Kleber, F. (2010). Vocal aging effects on F0 and the first formant: A longitudinal analysis in adult speakers. *Speech Communication*, *52*, 638–651.

Roberts, T., Morton, R., & Al-Ali, S. (2011). Microstructure of the vocal fold in elderly humans. *Clinical Anatomy*, *24*(5), 544–551. doi: 10.1002/ca.21114

Robinson, H., MacDonald, B., & Broadbent, E. (2014). The role of healthcare robots for older people at home: A review. *Int J of Soc Robotics*, *6*(4), 575–591. doi: 10.1007/s12369-014-0242-2

Rodrigues, C. (2016). Variação sociolinguística (Sociolinguistic variation). In A. M. Martins & E. Carrilho (Eds.), *Manual de linguística portuguesa.* De Gruyter. doi: https://doi.org/10.1515/9783110368840

Rohloff, M. (2020). Effects of f0 level, speech rate and articulation rate on age perception. In *SPPL2020: 2nd workshop on speech perception* (pp. 43–45). doi: 10.5281/zenodo.3732383

Ronneberger, O., Fischer, P., & Brox, T. (2015). U-Net: Convolutional networks for biomedical image segmentation. *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, *9351*, 234–241.

Roy, N., Nissen, S. L., Dromey, C., & Sapir, S. (2009). Articulatory changes in muscle tension dysphonia: Evidence of vowel space expansion following manual circumlaryngeal therapy. *Journal of Communication Disorders*, *42*(2), 124–135. doi: 10.1016/J.JCOMDIS.2008.10.001

Roy, N., Stemple, J., Merrill, R. M., & Thomas, L. (2007). Epidemiology of voice disorders in the elderly: Preliminary findings. *The Laryngoscope*, *117*(4), 628–633.

Rudnicka, E., Napierała, P., Podfigurna, A., Męczekalski, B., Smolarczyk, R., & Grymowicz, M. (2020). The World Health Organization (WHO) approach to healthy ageing. *Maturitas*, *139*, 6–11. doi: 10.1016/j.maturitas.2020.05.018

Russell, A., Penny, L., & Pemberton, C. (1995). Speaking fundamental frequency changes over time in women: A longitudinal study. *J Speech Lang Hear Res*, *38*(1), 101–109.

Ryan, W. J. (1972). Acoustic Aspects of the Aging Voice. *Journal of Gerontology*, *27*(2), 265–268. doi: 10.1093/geronj/27.2.265

Salema, L. (2007). *Consciência e caracterização fonética de variação dialectal em Português Europeu (Awareness and phonetic characterization of dialectal variation in European Portuguese)* (Master's thesis). Universidadde de Aveiro.

Salk, R. H., Hyde, J. S., & Abramson, L. Y. (2017). *Gender differences in depression in representative national samples: Meta-analyses of diagnoses and symptoms* (Vol. 143) (No. 8). American Psychological Association. doi: 10.1037/bul0000102

Santos, I. A. (1997). Fenómenos de palatalização vocálica na România: Valor dialectal em território português (continental). *Actas do XII encontro da APL*, 273–287.

Santos, I. A. (2003). *Variação linguística em espaço rural: a vogal [ü] numa comunidade do Baixo Mondego (Linguistic variation in rural areas: the vowel [ü] in a community from Baixo Mondego)*. Lisboa: Imprensa Nacional-Casa da Moeda.

Santos, I. R. (2005). *Análise Acústica da Voz de Indivíduos na Terceira Idade (Acoustic analysis the voice the elderly speakers)* (Master's thesis). Universidade de São Paulo, São Carlos.

Santos, M., Azevedo, S., Sousa, F., Machado, A. S., Santos, P. C., Freitas, S. V., ... da Silva, Á. M. (2021). Presbylarynx: Is it a sign of the health status of the elderly? *Journal of Voice*(in press). doi: 10.1016/J.JVOICE.2020.12.034

Sapir, S., Ramig, L. O., Spielman, J., & Fox, C. (2011). Acoustic metrics of vowel articulation in Parkinson's disease: Vowel space area (VSA) vs. vowel articulation index (VAI). In *MAVEBA-2011* (pp. 173–175). Firenze.

Sapir, S., Ramig, L. O., Spielman, J. L., & Fox, C. (2010). Formant Centralization Ratio: A proposal for a new acoustic measure of dysarthric speech. *J Speech Lang Hear Res*, *53*, 114–125.

Sarvasy, H., Elvin, J., Li, W., & Escudero, P. (2020). An acoustic phonetic description of Nungon vowels. *J. Acoust. Soc. Am.*, *147*(4), 2891–2900. doi: 10.1121/10.0001003

Sataloff, R. T. (1991). *Professional voice: The science and art of clinical care*. New York: Raven Press.

Sataloff, R. T. (2005). *Treatment of voice disorders*. San Diego: Plural Publishing.

Sataloff, R. T., Caputo Rosen, D., Hawkshaw, M., & Spiegel, J. R. (1997). The aging adult voice. *Journal of Voice*, *11*(2), 156–160. doi: 10.1016/s0892-1997(97)80072-0

Sataloff, R. T., Johns, M. M., & Kost, K. M. (2015). *Geriatric otolaryngology*. Stuttgart: Thieme Medical Publishers, Inc., and American Academy of Otolaryngology–Head and Neck Surgery Foundation.

Sataloff, R. T., & Kost, K. M. (2020). The effects of age on the voice, Part 1. *Journal of Singing*, *77*(1), 63–70.

Sataloff, R. T., Kost, K. M., & Linville, S. E. (2017). The effects of age on the voice. In R. T. Sataloff (Ed.), *Clinical assessment of voice* (2nd ed., pp. 221–240). San Diego: Plural Publishing, Inc.

Sato, K., & Hirano, M. (1996). Age-related changes of the macula flava of the human vocal fold. *Ann Otol Rhinol Laryngol*, *104*, 839–844. doi: 10.1177/000348949510401102

Sato, K., & Hirano, M. (1998). Age-related changes in the human laryngeal glands. *Ann Otol Rhinol*

Laryngol, *107*, 525–529.

Scherer, K. R. (1986). Vocal affect expression. A review and a model for future research. *Psychological Bulletin*, *99*(2), 143–165. doi: 10.1037/0033-2909.99.2.143

Scherer, S., Morency, L.-P., Gratch, J., & Pestian, J. (2015). Reduced vowel space is a robust indicator of psychological distress: A cross-corpus analysis. In *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)* (pp. 4789–4793). IEEE.

Scherer, S., Stratou, G., Mahmoud, M., Boberg, J., Gratch, J., Rizzo, A., & Morency, L.-P. (2013). Automatic behavior descriptors for psychological disorder analysis. In *International Conference and Workshops on Automatic Face and Gesture Recognition (FG)* (pp. 1–8). IEEE.

Schiel, F. (1999). Automatic phonetic transcription of non-prompted speech. In *14th ICPhS* (pp. 607–610). San Francisco.

Schietzel, S., Chocano-Bedoya, P. O., Sadlon, A., Gagesch, M., Willett, W. C., Orav, E. J., … Bischoff-Ferrari, H. A. (2022). Prevalence of healthy aging among community dwelling adults age 70 and older from five European countries. *BMC Geriatrics*, *22*(174), 1–11. doi: 10.1186/S12877-022-02755-8/TABLES/3

Schneider, S., Plank, C., Eysholdt, U., Schützenberger, A., & Rosanowski, F. (2011). Voice function and voice-related quality of life in the elderly. *Gerontology*, *57*(2), 109–114.

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

Schuller, B., Steidl, S., Batliner, A., Burkhardt, F., Devillers, L., Müller, C., & Narayanan, S. (2013). Paralinguistics in speech and language - State-of-the-art and the challenge. *Computer Speech and Language*, *27*, 4–39. doi: 10.1016/j.csl.2012.02.005

Schultz, B. G., Rojas, S., St John, M., Kefalianos, E., & Vogel, A. P. (2021). A cross-sectional study of perceptual and acoustic voice characteristics in healthy aging. *Journal of Voice*(in press). doi: 10.1016/j.jvoice.2021.06.007

Scobbie, J. M., Lawson, E., Cowen, S., Cleland, J., & Wrench, A. A. (2011). *A common co-ordinate system for mid-sagittal articulatory measurement.* Edinburgh.

Scukanec, G. P., Petrosino, L., & Squibb, K. (1991). Formant frequency characteristics of children, young adult, and aged female speakers. *Perceptual and motor skills*, *73*(1), 203–208.

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

Segura, L. (2013). Variedades dialetais do Português Europeu. In E. B. P. Raposo, M. F. B. do Nascimento, M. A. C. da Mota, L. Segura, & A. Mendes (Eds.), *Gramática do português volume 1* (pp. 85–136). Lisboa: Fundação Calouste Gulbenkian.

Selent, M. (2014). *Acoustic and respiratory measures as a function of age in the male voice* (Master's thesis). Cleveland State University.

Settersten, R. A., & Mayer, K. U. (1997). The measurement of age, age structuring, and the life course. *Annu. Rev. Sociol.*, *23*, 233–261. doi: 10.1146/ANNUREV.SOC.23.1.233

Sharma, G., & Goodwin, J. (2006). Effect of aging on respiratory system physiology and immunology. *Clinical Interventions in Aging*, *1*(3), 253–260. doi: 10.2147/CIIA.2006.1.3.253

Shrout, P. E., & Fleiss, J. L. (1979). Intraclass correlations: Uses in assessing rater reliability.

Psychological bulletin, *86*(2), 420–428.

Siegman, A. W., & Boyle, S. (1993). Voices of fear and anxiety and sadness and depression: The effects of speech rate and loudness on fear and anxiety and sadness and depression. *Journal of Abnormal Psychology*, *102*(3), 430–437. doi: 10.1037/0021-843X.102.3.430

Silva, D. J. (1994). The variable elision of unstressed vowels in European Portuguese: A case study. In S. C. Herring & J. C. Paolillo (Eds.), *UTA Working Papers in Linguistics* (Vol. 1, pp. 79–84). University of Texas at Arlington: Program in Linguistics.

Silva, D. J. (1998). Vowel lenition in São Miguel Portuguese. *Hispania*, *81*(1), 166–178.

Silva, P. T., Master, S., Andreoni, S., Pontes, P., Ramos, L. R., Da Silva, P. T., … Ramos, L. R. (2011, jul). Acoustic and Long-Term Average Spectrum measures to detect vocal aging in women. *Journal of Voice*, *25*(4), 411–419. doi: 10.1016/j.jvoice.2010.04.002

Smiljanic, R., & Gilbert, R. C. (2017). Acoustics of clear and noise-adapted speech in children, young, and older adults. *J Speech Lang Hear Res*, *60*(11), 3081–3096. doi: 10.1044/2017_JSLHR-S-16-0130

Smith, B. L., Wasowicz, J., & Preston, J. (1987). Temporal characteristics of the speech of normal elderly adults. *J Speech Hear Res*, *30*(4), 522–529.

Sobin, C., & Alpert, M. (1999). Emotion in speech: The acoustic attributes of fear, anger, sadness, and joy. *Journal of Psycholinguistic Research*, *28*(4), 347–365.

Sobin, C., & Sackeim, H. A. (1997). Psychomotor symptoms of depression. *Am J Psychiatry*, *154*(1), 4–17. doi: 10.1176/ajp.154.1.4

Song, J. Y. (2017). The use of ultrasound in the study of articulatory properties of vowels in clear speech. *Clin Linguist Phon*, *31*(5), 351–374. doi: 10.1080/02699206.2016.1268207

Sonies, B. C., Baum, B. J., & Shawker, T. H. (1984). Tongue motion in elderly adults: Initial in situ observations. *Journal of Gerontology*, *39*(3), 279–283. doi: 10.1093/geronj/39.3.279

Sonies, B. C., Stone, M., & Shawker, T. (1984). Speech and swallowing in the elderly. *Gerodontology*, *3*(2), 115–123. doi: 10.1111/j.1741-2358.1984.tb00362.x

Spazzapan, E. A., Cardoso, V. M., Fabron, E. M. G., Berti, L. C., Brasolotto, A. G., & Marino, V. C. d. C. (2018). Acoustic characteristics of healthy voices of adults: from young to middle age. *CoDAS*, *30*(5).

Spazzapan, E. A., Fabron, E. M. G., Berti, L. C., Chagas, E. F. B., Marino, V., Federighi Baisi Chagas, E., & Marino, V. C. d. C. (2020). Acoustic characteristics of the voice for Brazilian Portuguese speakers across the life span. *Journal of Voice*(In press). doi: 10.1016/j.jvoice.2020.09.019

Spazzapan, E. A., Marino, V. C. d. C., Cardoso, V. M., Berti, L. C., & Fabbron, E. M. G. (2019). Acoustic characteristics of voice in different cycles of life: An integrative literature review. *Rev. CEFAC*, *21*(3), e15018. doi: 10.1590/1982-0216/201921315018

Sperry, E. E., & Klich, R. J. (1992). Speech breathing in senescent and younger women during oral reading. *J Speech Hear Res*, *35*(6), 1246–1255.

Staiger, A., Schölderle, T., Brendel, B., & Ziegler, W. (2017). Dissociating oral motor capabilities: Evidence from patients with movement disorders. *Neuropsychologia*, *95*, 40–53. doi: 10.1016/j.neuropsychologia.2016.12.010

Stassen, H. H., Bomben, G., & Günther, E. (1991). Speech characteristics in depression. *Psychopathol-*

ogy, *24*(2), 88–105.

Stassen, H. H., Kuny, S., & Hell, D. (1998). The speech analysis approach to determining onset of improvement under antidepressants. *European Neuropsychopharmacology*, *8*(4), 303–310. doi: 10.1016/S0924-977X(97)00090-4

Stathopoulos, E. T., Huber, J. E., & Sussman, J. E. (2011). Changes in acoustic characteristics of the voice across the life span: Measures from individuals 4–93 years of age. *J Speech Lang Hear Res*, *54*(4), 1011–1021.

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*.

Statistics Portugal. (2021). CENSOS 2021 – Resultados provisórios (CENSOS 2021 - Preliminary results). *Destaque: informação à comunicação social*.

Steffens, Y. (2011). *The aging voice*. GRIN Verlag.

Stoicheff, M. L. (1981). Speaking fundamental frequency characteristics of nonsmoking female adults. *J Speech Hear Res*, *24*, 437–441. doi: 10.1044/jshr.2403.437

Stone, M. (2005). A guide to analysing tongue motion from ultrasound images. *Clin Linguist Phon*, *19*(6-7), 455–501. doi: 10.1080/02699200500113558

Strycharczuk, P., & Scobbie, J. M. (2015). Velocity measures in ultrasound data. Gestural timing of post-vocalic/l/in English. In *8th ICPhS.* Glasgow.

Strycharczuk, P., & Scobbie, J. M. (2017). Fronting of Southern British English high-back vowels in articulation and acoustics. *J. Acoust. Soc. Am.*, *142*(1), 322–331.

Styler, W. (2015). *On the acoustical and perceptual features of vowel nasality* (Doctoral dissertation). University of Colorado.

Tabain, M. (2013). Research methods in speech production. In M. J. Jones & R.-A. Knight (Eds.), *Bloomsbury companion to phonetics* (pp. 39–56). London: Bloomsbury.

Taguchi, T., Tachikawa, H., Nemoto, K., Suzuki, M., Nagano, T., Tachibana, R., … Arai, T. (2018). Major depressive disorder discrimination using vocal acoustic features. *Journal of affective disorders*, *225*, 214–220.

Takano, S., Kimura, M., Nito, T., Imagawa, H., Sakakibara, K.-I., & Tayama, N. (2010). Clinical analysis of presbylarynx—Vocal fold atrophy in elderly individuals. *Auris Nasus Larynx*, *37*(4), 461–464. doi: 10.1016/j.anl.2009.11.013

Tarafder, K. H., Datta, P. G., & Tariq, A. (2012). The aging voice. *Bangabandhu Sheikh Mujib Medical University Journal*, *5*(1), 83–86.

Teixeira, A., Braga, D., Coelho, L., Fonseca, J. A., Alvarelhão, J., Martín, I., … Dias, M. (2009). Speech as the basic interface for assistive technology. In *DSAI'2009.*

Teixeira, V., Pires, C., Pinto, F., Freitas, J., Dias, M. S., & Rodrigues, E. M. (2012). Towards elderly social integration using a multimodal human-computer interface. In *International Living Usability Lab Workshop on AAL Latest Solutions, Trends and Applications.*

The MathWorks. (2021). *Choose classifier options - MATLAB & Simulink.* Retrieved 2021-01-15, from `https://www.mathworks.com/help/stats/choose-a-classifier.html`

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

Thomas, L. B., Harrison, A. L., & Stemple, J. C. (2008, jul). Aging thyroarytenoid and limb skeletal muscle: Lessons in contrast. *Journal of Voice*, *22*(4), 430–450. doi: 10.1016/J.JVOICE.2006.11.006

Thompson, L. A., & Chen, H. (2021). Physiology of aging of older adults: Systemic and oral health considerations — 2021 Update. *Dental Clinics of North America*, *65*(2), 275–284. doi: 10.1016/J.CDEN.2020.11.002

Titze, I. R. (1994). Voice classification and life-span changes. In *Principles of voice production* (pp. 169–190). Englewood Cliffs, New Jersey: Prentice Hall.

Tolkmitt, F., Helfrich, H., Standke, R., & Scherer, K. (1982). Vocal indicators of psychiatric treatment effects in depressives and schizophrenics. *Journal of Communication Disorders*, *15*(3), 209–222. doi: 10.1016/0021-9924(82)90034-X

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

Trevino, A. C., Quatieri, T. F., & Malyska, N. (2011). Phonologically-based biomarkers for major depressive disorder. *EURASIP Journal on Advances in Signal Processing*, *2011*(1), 42.

Troen, B. R. (2003). The biology of aging. *Mount Sinai Journal of Medicine*, *70*(1), 3–22.

Tuomainen, O., Hazan, V., & Taschenberger, L. (2019). Speech communication in background noise: Effects of aging. In *19th ICPhS* (pp. 805–809). Australasian Speech Science and Technology Association Inc.

Turgeon, C., Trudeau-Fisette, P., Fitzpatrick, E., & Ménard, L. (2017). Vowel intelligibility in children with cochlear implants: An acoustic and articulatory study. *International Journal of Pediatric Otorhinolaryngology*, *101*, 87–96. doi: 10.1016/J.IJPORL.2017.07.022

Turk, L. M., & Hogg, D. A. (1993). Age changes in the human laryngeal cartilages. *Clinical Anatomy*, *6*, 154–162. doi: 10.1002/CA.980060305

Türkmen, S., Cansu, A., Türedi, S., Eryigit, U., Sahin, A., Gündüz, A., & Shavit, I. (2012, nov). Age-dependent structural and radiological changes in the larynx. *Clinical Radiology*, *67*(11), e22–e26. doi: 10.1016/J.CRAD.2012.07.006

Tykalova, T., Skrabal, D., Boril, T., Cmejla, R., Volin, J., & Rusz, J. (2020). Effect of ageing on acoustic characteristics of voice pitch and formants in Czech vowels. *Journal of Voice*. doi: 10.1016/j.jvoice.2020.02.022

United Nations, Department of Economic and Social Affairs, Population Division. (2019). *World Population Ageing 2019: Highlights.* (ST/ESA/SER.A/430).

Valente, A. R., Oliveira, C., Albuquerque, L., Teixeira, A., & Barbosa, P. A. (2021). Prosodic changes with age: A longitudinal study on a famous European Portuguese native speaker. *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, *12997 LNAI*, 726–736. doi: 10.1007/978-3-030-87802-3_65

Van de Velde, S., Bracke, P., & Levecque, K. (2010). Gender differences in depression in 23 European countries. Cross-national variation in the gender gap in depression. *Social Science & Medicine*, *71*(2), 305–313. doi: 10.1016/J.SOCSCIMED.2010.03.035

Van der Harst, S., & 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. doi: 10.1017/S0954394514000040

Vandervoort, A. A. (2002). Aging of the human neuromuscular system. *Muscle & Nerve*, *25*(1), 17–25. doi: 10.1002/MUS.1215

Vanderwegen, J., Guns, C., Van Nuffelen, G., Elen, R., & De Bodt, M. (2013). The influence of age, sex, bulb position, visual feedback, and the order of testing on maximum anterior and posterior tongue strength and endurance in healthy Belgian adults. *Dysphagia*, *28*, 159–166. doi: 10.1007/S00455-012-9425-X

Veloso, J. (2007). Schwa in European Portuguese: The phonological status of [ɨ]. In *Journees d'Etudes Linguistiques* (pp. 55–60). Nantes.

Veloso, J. (2010). Central, epenthetic, unmarked vowels and schwas: A brief outline of some essential differences. *Linguística: Revista de Estudos Linguísticos da Universidade do Porto*, *5*, 191–213.

Verdonck-de Leeuw, I. M., & Mahieu, H. F. (2004). Vocal aging and the impact on daily life: A longitudinal study. *Journal of Voice*, *18*(2), 193–202.

Verhoeven, J., De Pauw, G., & Kloots, H. (2004). Speech rate in a pluricentric language: A comparison between Dutch in Belgium and the Netherlands. *Language and Speech*, *43*(3), 297–308.

Ververidis, D., & Kotropoulos, C. (2006). Emotional speech recognition: Resources, features, and methods. *Speech Communication*, *48*, 1162–1181.

Vicsi, K., Sztahó, D., & Kiss, G. (2012). Examination of the sensitivity of acoustic-phonetic parameters of speech to depression. In *2012 IEEE 3rd International Conference on Cognitive Infocommunications (CogInfoCom)* (pp. 511–515). IEEE.

Vigário, M., & Frota, S. (2003). The intonation of Standard and Northern European Portuguese: A comparative intonational phonology approach. *Journal of Portuguese Linguistics*, *2*(2), 115. doi: 10.5334/JPL.31/

Vipperla, R. (2011). *Automatic speech recognition for ageing voices* (Doctoral dissertation). University of Edinburgh.

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

Volín, J., Tykalová, T., & Bořil, T. (2017). Stability of prosodic characteristics across age and gender groups. In *INTERSPEECH* (pp. 3902–3906). Stockholm. doi: 10.21437/Interspeech.2017-1503

Waller, S. S. (2019). *Estimation of speaker age: Effects of speech properties and speech material* (Doctoral dissertation). Mid Sweden University.

Watson, C. I., & Evans, Z. E. (2016). Sound change or experimental artifact?: A study on the impact of data preparation on measuring sound change. In *16th Australasian International Conference on Speech Science and Technology* (pp. 261–264). Parramatta, Australia.

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

Watson, C. I., & Hui, C. T. J. (2010). Two short studies in vocal tract measurements. In *13th Australasian International Conference on Speech Science and Technology.* Melbourne: The

Australasian Speech Science & Technology Association.

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

Weeks, J. W., Lee, C.-Y., Reilly, A. R., Howell, A. N., France, C., Kowalsky, J. M., & Bush, A. (2012). "The Sound of Fear": Assessing vocal fundamental frequency as a physiological indicator of social anxiety disorder. *Journal of anxiety disorders*, *26*(8), 811–822.

Whalen, D. H., & Levitt, A. G. (1995). The universality of intrinsic F0 of vowels. *Journal of Phonetics*, *23*(3), 349–366.

Wieling, M., Tomaschek, F., Arnold, D., Tiede, M., Bröker, F., Thiele, S., ... Baayen, R. H. (2016). Investigating dialectal differences using articulography. *Journal of Phonetics*, *59*, 122–143. doi: https://doi.org/10.1016/j.wocn.2016.09.004

Willcox, D. C., Willcox, B. J., & Poon, L. W. (2010). Centenarian studies: Important contributors to our understanding of the aging process and longevity. *Current Gerontology and Geriatrics Research*, *2010*. doi: 10.1155/2010/484529

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

Winkelmann, R., Jaensch, K., Cassidy, S., & Harrington, J. (2020). *Main package of the EMU speech database management system.* Retrieved from `https://github.com/IPS-LMU/emuR`

Winkler, R. (2004). Open Quotient and breathiness in aging voices-changes with increasing chronological age and it's perception [Poster]. In *German-french summerschool on "cognitive and physical models of speech production, perception and perception-production interaction"*. Lubmin, Germany.

Wohlert, A. B., & Smith, A. (1998). Spatiotemporal stability of lip movements in older adult speakers. *J Speech Lang Hear Res*, *41*(1), 41–50. doi: 10.1044/jslhr.4101.41

Won, E., & Kim, Y.-K. (2016). Stress, the autonomic nervous system, and the immune-kynurenine pathway in the etiology of depression. *Current neuropharmacology*, *14*(7), 665–673. doi: 10.2174/1570159X14666151208113006

Woo, P., Casper, J., Colton, R., & Brewer, D. (1992). Dysphonia in the aging: Physiology versus disease. *Laryngoscope*, *102*(2), 139–144. doi: 10.1288/00005537-199202000-00007

World Health Organization. (2007). *Women, ageing and health: A framework for action - Focus on gender*. Geneva: WHO publications.

World Health Organization. (2017a). *Mental health of older adults.* Retrieved 2022-05-21, from `https://www.who.int/news-room/fact-sheets/detail/mental-health-of-older-adults`

World Health Organization. (2017b). *WHO Clinical Consortium on Healthy Ageing. Topic focus: frailty and intrinsic capacity*. Geneva, Switzerland: Report of consortium meeting 1–2 December 2016 (WHO/FWC/A).

World Health Organization. (2018). *Mental disorders.* Retrieved 2019-09-21, from `https://www.who.int/en/news-room/fact-sheets/detail/mental-disorders`

World Health Organization. (2020). *Decade of Healthy Ageing 2020-2030* (Tech. Rep.). Retrieved from `https://www.who.int/initiatives/decade-of-healthy-ageing`

Wörtwein, T., Morency, L., & Scherer, S. (2015). Automatic assessment and analysis of public speaking anxiety: A virtual audience case study. In *International Conference on Affective Computing and Intelligent Interaction (ACII)* (pp. 187–193). doi: 10.1109/ACII.2015.7344570

Wrench, A. A., & Scobbie, J. M. (2008). High-speed cineloop ultrasound vs. video ultrasound tongue imaging: Comparison of front and back lingual gesture location and relative timing. In *8th International Seminar on Speech Production (ISSP).*

Ximenes Filho, J. A., Tsuji, D. H., Nascimento, P. H. S. D., & Sennes, L. U. (2003, oct). Histologic changes in human vocal folds correlated with aging: A histomorphometric study. *Annals of Otology, Rhinology and Laryngology*, *112*(10), 894–898. doi: 10.1177/000348940311201012

Xue, S. A., & Deliyski, D. (2001). Effects of aging on selected acoustic voice parameters: Preliminary normative data and educational implications. *Educational Gerontology*, *27*(2), 159–168.

Xue, S. A., & 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. doi: 10.1044/1092-4388(2003/054)

Xue, S. A., Jiang, J., Lin, E., Glassenberg, R., & Mueller, P. B. (1999). Age-related changes in human vocal tract configurations and the effects on speakers' vowel formant frequencies: A pilot study. *Logopedics Phoniatrics Vocology*, *24*(3), 132–137. doi: 10.1080/140154399435084

Yang, Y., Fairbairn, C., & Cohn, J. F. (2013). Detecting depression severity from vocal prosody. *IEEE Transactions on Affective Computing*, *4*(2), 142–150.

Yuan, J. (2013). The spectral dynamics of vowels in Mandarin Chinese. In *INTERSPEECH* (pp. 1193–1197). Lyon, France.

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

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

Zamponi, V., Mazzilli, R., Mazzilli, F., & Fantini, M. (2021). Effect of sex hormones on human voice physiology: From childhood to senescence. *Hormones*, *20*(4), 691–696. doi: 10.1007/S42000-021-00298-Y/TABLES/2

Zharkova, N. (2013). A normative-speaker validation study of two indices developed to quantify tongue dorsum activity from midsagittal tongue shapes. *Clinical Linguistics & Phonetics*, *27*(6-7), 484–496. doi: 10.3109/02699206.2013.778903

Zharkova, N., Hewlett, N., & Hardcastle, W. J. (2008). Analysing coarticulation in Scottish English children and adults: An ultrasound study. *Canadian Acoustics*, *36*(3), 158–159.

Zharkova, N., Hewlett, N., & Hardcastle, W. J. (2011). Coarticulation as an indicator of speech motor control development in children: An ultrasound study. *Motor Control*, *15*(1), 118–140.

Zharkova, N., Hewlett, N., & Hardcastle, W. J. (2012). An ultrasound study of lingual coarticulation in/s V/syllables produced by adults and typically developing children. *JIPA*, *42*(2), 193–208.

Zharkova, N., Lickley, R., & Hardcastle, W. J. (2014). Development of lingual coarticulation and articulatory constraints between childhood and adolescence: an ultrasound study. In S. Fuchs,

M. Grice, A. Hermes, L. Lancia, & D. Mücke (Eds.), *10th International Seminar on Speech Production (ISSP)* (pp. 472–475). Cologne.

Zhu, J., Styler, W., & Calloway, I. (2019). A CNN-based tool for automatic tongue contour tracking in ultrasound images. *eprint arXiv:1907.10210*, 1–6.

Zigmond, A. S., & Snaith, R. P. (1983). The hospital anxiety and depression scale. *Acta psychiatrica scandinavica*, *67*(6), 361–370.

ETHICAL APPROVAL (ACOUSTIC DATABASE)

This page intentionally contains only this sentence.

Parecer da Comissão de Ética para a Saúde do
Centro Hospitalar de São João / Faculdade de Medicina da Universidade do Porto

Título do Projecto: Envelhecimento vocal: estudo acústico das alterações de fala com a idade

Nome da Investigadora Principal: Dra. Luciana Albuquerque, Terapeuta da Fala, aluna no Programa Doutoral em Gerontologia e Geriatria (afiliado à FMUP – CINTESIS)

Onde decorre o Estudo: Universidades seniores, e instituições afins

Objectivos do Estudo:
Este trabalho de investigação tem como principal objectivo investigar acusticamente as vogais produzidas por indivíduos saudáveis do Português Europeu (PE), divididos em diferentes faixas etárias, de modo a averiguar o efeito da idade e do género nas características acústicas da fala; estudar o impacto das alterações acústicas relacionadas com a idade na qualidade de vida; analisar a relação entre variáveis de bem-estar psicológico e as características acústicas das vogais.

Estudo realizado no âmbito do Programa Doutoral em Gerontologia e Geriatria (afiliado à FMUP – CINTESIS), sob orientação da Prof.ª Doutora Catarina Alexandra Monteiro de Oliveira e co-orientação do Prof. Doutor António Joaquim da Silva Teixeira e da Prof.ª Doutora Daniela Maria Pias de Figueiredo, cujas declarações anexa.

Concepção e Pertinência do estudo:
Ao contrário de outras línguas, para o Português Europeu (PE) praticamente não existem dados sobre os correlatos acústicos do envelhecimento. Além disso, pouco se sabe sobre a autoperceção das alterações vocais com a idade, assim como, sobre o efeito destas alterações sobre na qualidade de vida dos idosos. Estas informações serão importantes, tanto para a prática clínica dos Terapeutas da Fala, como para o desenvolvimento de tecnologias de fala (sintese e reconhecimento) para o PE, mais adaptadas às necessidades do utilizador final.
Deste modo e para responder aos objetivos propostos, a investigadora pretende gravar uma amostra de fala (PE) de ambos os géneros, para posteriormente serem analisados acusticamente. Estão claramente definidos os critérios de inclusão e exclusão no estudo.
Faz também parte da recolha de dados a utilização de questionário sociodemográfico e duas escalas (qualidade de vida e voz e Ansiedade e Depressão Hospitalar em anexo).
Benefício/risco:
A participação no estudo não acarreta riscos, uma vez que só será gravada uma amostra da voz.
Confidencialidade dos dados:
Será utilizado um código numérico para identificação de cada participante.

Respeito pela liberdade e autonomia do sujeito de ensaio: Salvaguardado pelo consentimento informado, esclarecido e livre, da CES e adequada informação ao participante

***Curriculum* da investigadora**: Adequado à investigação.

Data previsível da conclusão do estudo: Julho de 2019

Conclusão: Proponho um parecer favorável à realização deste projecto de investigação.

Porto, 22 de fevereiro de 2018

A Relatora da CES, Enf.ª Teresa Guerreiro

Comissão de Ética Centro Hospitalar São João /
/ Faculdade de Medicina da Universidade do Porto n.º 38 / 18

SÃO JOÃO

U.PORTO
FMUP FACULDADE DE MEDICINA UNIVERSIDADE DO PORTO

Questionário para submissão de Investigação

Exmo. Sr. Presidente da Comissão de Ética do Centro Hospitalar de São João/ Faculdade de Medicina da Universidade do Porto,

Pretendendo realizar a investigação infracitada, solicito a V. Exa., na qualidade de Investigador, a sua apreciação e a elaboração do respetivo parecer. Para o efeito, anexo toda a documentação requerida.

IDENTIFICAÇÃO DO ESTUDO

Título da investigação: Envelhecimento Vocal: Estudo acústico das alterações de fala com a idade

Nome do investigador: Luciana Patrícia Martins Nunes Pereira Albuqerque

Endereço eletrónico: lucianapereira@ua.pt **Contacto telefónico:** 912758447

Caracterização da investigação:

☐ Estudo retrospetivo ☒ Estudo observacional ☐ Estudo prospetivo

☐ Inquérito ☐ Outro. Qual? ______

Tipo de investigação:

☐ Com intervenção ☒ Sem intervenção

Formação do investigador em boas práticas clínicas (*GCP*): ☐ Sim ☒ Não

Promotor *(se aplicável)*: ______

Nome do orientador de dissertação/tese *(se aplicável)*: Professor(a) Catarina Oliveira / António Teixeira/ Daniela Figueiredo

Endereço eletrónico: ______

Local/locais onde se realiza a investigação: Universidade de Aveiro, Universidades Seniores e instituições afins

Data prevista para início: 01 / 02 / 2018 Data prevista para o término: 31 / 07 / 2019

PROTOCOLO DO ESTUDO

Síntese dos objetivos:

1º - Analisar acusticamente as vogais produzidas por indivíduos saudáveis do Português Europeu (PE), divididos em diferentes faixas etárias, de modo a averiguar o efeito da idade e do género nas características acústicas da fala;
2º - Estudar o impacto das alterações acústicas relacionadas com a idade na qualidade de vida;
3º - Analisar a relação entre variáveis de bem-estar psicológico e as características acústicas das vogais.

Fundamentação ética (ganhos em conhecimento/inovação; ponderação benefícios/riscos):

Este estudo é importante na medida em que para o PE praticamente não existem dados sobre os correlatos acústicos do envelhecimento. Para além disso, pouco se sabe sobre a autoperceção das alterações vocais com a idade, assim como, sobre o efeito destas alterações na qualidade de vida dos idosos. Esta informações serão importantes tanto para a prática clínica dos Terapeutas da Fala, como para o desenvolvimento de tecnologias de fala (síntese e reconhecimento) para o PE, mais adaptadas às necessidades do utilizador final.
Este trabalho não implicará riscos (físicos, emocionais e financeiros) significativamente superiores aos do dia-a-dia, tanto para os participantes como para o investigador, dado que para a realização do estudo só será necessária a gravação de uma amostra de fala. Será garantida a confidencialidade dos dados e de todas as informações recolhidas junto dos participantes, sendo apenas alvo de processamento os dados acústicos e os dados referentes ao género, idade, qualidade de vida e bem-estar.
Este trabalho não traz benefícios diretos para os participantes, no entanto, permitirá certamente compreender melhor o processo de envelhecimento vocal, assim como estudar as alterações vocais em função do bem-estar psicológico.

CONFIDENCIALIDADE

De que forma é garantida a anonimização dos dados recolhidos de toda a informação?

Será utilizado um código numérico para identificação de cada um dos participantes.

O investigador necessita ter acesso a dados do processo clínico? ☐ Sim ☒ Não

Está previsto o registo de imagem ou som dos participantes? ☒ Sim ☐ Não

Se sim, está prevista a destruição deste registo após o sua utilização? ☐ Sim ☒ Não

CONSENTI MENTO

O estudo implica recrutamento de:

Doentes: ☐ Sim ☒ Não Voluntários saudáveis: ☒ Sim ☐ Não

Menores de 18 anos: ☐ Sim ☒ Não

Outras pessoas sem capacidade do exercício de autonomia: ☐ Sim ☒ Não

A investigação prevê a obtenção de Consentimento Informado: ☒ Sim ☐ Não

Se não, referir qual o fundamento para a isenção:

Existe informação escrita aos participantes: ☒ Sim ☐ Não

PROPRIEDADE DOS DADOS

A investigação e os seus resultados são propriedade intelectual de:

☒ Investigador ☐ Promotor ☐ Ambos ☐ Serviço onde é realizado

☐ Não aplicável Outro: ____

BENEFÍCIOS, RISCOS E CONTRAPARTIDAS PARA OS PARTICIPANTES

Benefícios previsíveis:

Este trabalho não traz benefícios diretos para os participantes.

Riscos/incómodos previsíveis:

Esta investigação não implicará riscos (físicos, emocionais e/ou financeiros) tanto para os participantes como para o investigador.

São dadas contrapartidas aos participantes:

- *pela participação* ☐ Sim ☒ Não ☐ Não aplicável
- *pelas deslocações* ☐ Sim ☒ Não ☐ Não aplicável
- *pelas faltas ao emprego* ☐ Sim ☒ Não ☐ Não aplicável
- *por outras perdas e danos* ☐ Sim ☒ Não ☐ Não aplicável

CUSTOS / PLANO FINANCEIRO

Os custos da investigação são suportados por:

☐ Investigador ☐ Promotor ☐ Serviço onde é realizado

☐ Não aplicável Outro: Fundação para a Ciência e a Tecnologia (FCT) (bolsa com a referência SFRH/BD/115381/2016)

LISTA DE DOCUMENTOS ANEXOS

- [ ] Pedido de autorização ao Presidente do Conselho de Administração do Centro Hospitalar de São João *(se aplicável)*
- [ ] Pedido de autorização à Diretora da Faculdade de Medicina da Universidade do Porto *(se aplicável)*
- [x] Protocolo do estudo
- [ ] Declaração do Diretor de Serviço onde decorre o estudo
  (sendo um estudo na área de enfermagem deve anexar também a concordância da chefia de enfermagem)
- [ ] Profissional de ligação
- [x] Informação dos orientadores
- [x] Informação ao participante
- [x] Modelo de consentimento
- [x] Instrumentos a utilizar *(inquéritos, questionários, escalas, p.ex.)*: Questionário sociodemográfico; HADS; QVV-Luso; Corpus de fala
- [x] Curriculum Vitae abreviado *(máx. 3 páginas)*
- [ ] Protocolo financeiro
- [ ] Outros:

COMPROMISSO DE HONRA E DECLARAÇÃO DE INTERESSES

Declaro por minha honra que as informações prestadas neste questionário são verdadeiras. Mais declaro que, durante o estudo, serão respeitadas as recomendações constantes da Declaração de Helsínquia (1960 e respetivas emendas), e da Organização Mundial da Saúde, Convenção de Oviedo e das "Boas Práticas Clínicas" (GCP/ICH) no que se refere à experimentação que envolve seres humanos. Aceito, também, a recomendação da CES de que o recrutamento para este estudo se fará junto de doentes que não tenham participado em outro estudo, nos últimos três meses. Comprometo-me a entregar à CES o relatório final da investigação, assim que concluído.

Porto, 3 *de* janeiro *de* 2018

Nome legível: Luciana Patrícia Martins Nunes Pereira Albuquerque

Luciana Albuquerque
assinatura

Parecer da Comissão de Ética do Centro Hospitalar de São João/ FMUP Emitido na reunião plenária da CE de 16 / 02 / 2018

A Comissão de Ética para a Saúde tendo aprovado o parecer do Relator, aguarda que o Investigador/Promotor esclareça as questões nele enunciadas para que possa emitir parecer definitivo.

Centro Hospitalar **São João**.

CONSIDERADOS QUE FORAM COMO SATISFATÓRIOS OS ESCLARECIMENTOS PRESTADOS PELO(A) INVESTIGADOR(A), A CES APROVA POR UNANIMIDADE O PARECER DO RELATOR, PELO QUE NADA TEM A OPOR À REALIZAÇÃO DESTE PROJETO DE INVESTIGAÇÃO.

22/02/2018

This page intentionally contains only this sentence.

CONSENT FORM (ACOUSTIC DATABASE)

This page intentionally contains only this sentence.

Envelhecimento Vocal: Estudo acústico das alterações de fala coma a idade

Folha de Informação aos Participantes

Eu, Luciana Albuquerque, Terapeuta da Fala (Cédula Profissional n.º C-038616181), a realizar o Doutoramento em Gerontologia e Geriatria (Programa Doutoral conjunto entre Universidade de Aveiro/Universidade do Porto), sob orientação da Professora Doutora Catarina Oliveira (ESSUA/IEETA) e coorientação do Professor Doutor António Teixeira (DETI/IEETA) e da Professora Doutora Daniela Figueiredo (ESSUA/CINTESIS-UA), venho solicitar a Sua colaboração no estudo intitulado "Envelhecimento Vocal: Estudo acústico das alterações de fala com a idade".

O principal objetivo deste estudo é analisar as características acústicas das vogais do Português Europeu (PE), em função da idade, do género e da qualidade de vida.

A sua participação neste estudo envolverá a gravação de uma amostra de fala, tendo em vista a obtenção de parâmetros acústicos (ex. frequência fundamental, formantes e duração das vogais) que, posteriormente, serão sujeitos a uma análise estatística.

Ser-lhe-á pedido que produza, num tom e ritmo de fala que considere normal, um conjunto de palavras, que serão repetidas 3 vezes. Estas serão apresentadas num ecrã de computador com recurso a imagem e representação ortográfica. Adicionalmente, terá de descrever uma imagem. A recolha dos dados será realizada utilizando um microfone conectado a uma placa de som externa. Ser-lhe-á ainda solicitado que preencha um pequeno questionário de caracterização sociodemográfica e de autoperceção da qualidade de vida e bem-estar psicológico. Todos os procedimentos terão uma duração aproximada de 30 minutos.

Os resultados do estudo poderão vir a ser divulgados em revistas científicas e/ou em congressos/eventos da área.

Será garantida a confidencialidade e anonimato de todas as informações recolhidas, sendo apenas alvo de processamento os dados acústicos e os dados referentes ao género, idade, qualidade de vida e bem-estar. De modo a garantir o anonimato, será utilizado um código numérico para identificação de cada um dos participantes, atribuído no momento de recolha da informação pessoal.

A sua participação é voluntária, pelo que poderá optar, a qualquer momento, por desistir do estudo, sem que daí advenham quaisquer prejuízos ou consequências. Este estudo não implicará qualquer risco físico e/ou emocional e não envolverá contrapartidas financeiras, tendo sido alvo de análise e aprovação pela Comissão de Ética Centro Hospitalar de S. João / FMUP.

A equipa de investigadores está ao seu dispor para responder a qualquer dúvida que considere pertinente.

Contacto da investigadora responsável (caso queira colocar dúvidas ou questões):
Telemóvel: 912758447 / e-mail: lucianapereira@ua.pt

CONSENTIMENTO INFORMADO, ESCLARECIDO E LIVRE

SÃO JOÃO

PARA INVESTIGAÇÃO CLÍNICA

Considerando a "Declaração de Helsínquia" da Associação Médica Mundial (Helsínquia 1964; Tóquio 1975; Veneza 1983; Hong Kong 1989; Somerset West 1996, Edimburgo 2000, Seoul 2008, Fortaleza 2013)

Designação do Estudo *(em português)*

Confirmo que expliquei ao participante/ representante legal, de forma adequada e compreensível, a investigação referida, os benefícios, os riscos e possíveis complicações associadas à sua realização.

Informação escrita em anexo: ☐ Não ☐ Sim (Nº de páginas________)

O Investigador responsável

Nome: ______________________________ ____________________

legível assinatura

Identificação do participante

Nome: ____

BI/ CC nº: ______________

Participante/ Representante legal

· Compreendi a explicação que me foi facultada acerca do estudo que se tenciona realizar: os objetivos, os métodos, os benefícios previstos, os riscos potenciais e o eventual desconforto.

· Solicitei todas as informações de que necessitei, sabendo que o esclarecimento é fundamental para uma boa decisão.

· Fui informado da possibilidade de livremente recusar ou abandonar a todo o tempo a participação no estudo, sem que isso possa ter como efeito qualquer prejuízo na assistência que é prestada.

·Declaro não ter sido incluído em nenhum outro projeto de investigação nos últimos três meses.

Concordo com a participação neste estudo, de acordo com os esclarecimentos que me foram prestados, como consta neste documento, do qual me foi entregue uma cópia.

Data: ____/____/____ ____________________

assinatura

Nome *(Pais/Representante legal)*: ______________________________

BI/ CC nº: ______________ Grau de parentesco: ____________________

Data: ____/____/____ ____________________

assinatura

BACKGROUND QUESTIONNAIRE (ACOUSTIC DATABASE)

This page intentionally contains only this sentence.

Envelhecimento Vocal: Estudo acústicos das alterações de fala com a idade

Eu, Luciana Albuquerque, Terapeuta da Fala (Cédula Profissional n.º C-038616181), a realizar o Doutoramento em Gerontologia e Geriatria, sob orientação da Professora Doutora Catarina Oliveira (ESSUA/IEETA) e coorientação do Professor Doutor António Teixeira (DETI/IEETA) e da Professora Doutora Daniela Figueiredo (ESSUA/CINTESIS-UA), venho solicitar a Sua colaboração no presente estudo. O principal objetivo deste estudo é analisar as características acústicas das vogais do Português Europeu (PE), em função da idade, do género e da qualidade de vida. Para tal, agradeço que preencha o questionário sociodemográfico que se segue, que tem como objetivo recolher informações relacionadas com as suas características vocais.

Será garantida a confidencialidade e anonimato de todas as informações recolhidas, sendo estas apenas utilizadas para efeitos de análise estatística. A sua participação é voluntária, pelo que poderá optar, a qualquer momento, por desistir do estudo, sem que daí advenham quaisquer prejuízos ou consequências. A equipa de investigadores está ao seu dispor para responder a qualquer dúvida que considere pertinente. Muito obrigada pela sua disponibilidade!

*Obrigatório

1. **Código do Participante** *

   Fornecido pela investigadora

   ____

2. **Data de Nascimento** *

   ____

   Exemplo: 15 de dezembro 2012

3. **Sexo** *

   Marcar apenas uma oval.

   - ◯ Feminino
   - ◯ Masculino

4. **Habilitações académicas** *

   Marcar apenas uma oval.

   - ◯ Não frequentou a escola ou não concluiu o 1.º Ciclo
   - ◯ 1.º Ciclo (4.º ano de escolaridade)
   - ◯ 2.º Ciclo (6.º ano de escolaridade)
   - ◯ 3.º Ciclo ou formação equivalente (9.º ano de escolaridade)
   - ◯ Ensino secundário ou formação equivalente (12.º ano de escolaridade)
   - ◯ Formação superior profissional
   - ◯ Bacharelato
   - ◯ Licenciatura
   - ◯ Mestrado
   - ◯ Doutoramento
   - ◯ Outra: ______________________________

5 **Situação Profissional** *
Indique a profissão que desempenha atualmente ou a profissão que desempenhou durante um maior número de anos

6. **Nasceu em que concelho** *

7. **Vive em que concelho** *

8. **Em alguma fase da sua vida viveu fora do país?** *
Marcar apenas uma oval.

- Não *Passe para a pergunta 11.*
- Sim *Passe para a pergunta 9.*

Período de ausência do país

9. **Durante quanto tempo esteve fora do país** *

10. **Onde esteve durante esse período** *

Informação sobre saúde

11. **Como classifica o seu estado de saúde geral?** *
Marcar apenas uma oval.

- Mau
- Médio
- Bom
- Excelente

12 **Tem ou já teve algum (a):** *

Pode assinalar mais do que uma opção:
Marcar tudo o que for aplicável.

- [ ] Traumatismo Cranioencefálico
- [ ] Acidente Vascular Cerebral (AVC)
- [ ] Doença Neurológica (ex. Esclerose Múltipla, Epilepsia, Parkinson)
- [ ] Perturbação Neurocognitiva (ex. Demência)
- [ ] Alteração Emocional (ex. depressão, fadiga crónica, estados de ansiedade, insónias)
- [ ] Cancro/Tumor de cabeça ou pescoço (ex. tiroide, garganta)
- [ ] Problema hormonal/ endócrino (ex. hipertiroidismo, hipotiroidismo)
- [ ] Azia crónica ou refluxo
- [ ] Perturbação de Fala/ Voz ou Linguagem (ex. gaguez, laringite crónica, paralisia das pregas vocais, afasia)
- [ ] Problema Respiratório crónico (ex. asma, rinite, sinusite)
- [ ] Cirurgia oral ou facial
- [ ] Cirurgia de cabeça ou pescoço
- [ ] Cirurgia do tórax
- [ ] Não

13. **Teve alguma infeção respiratória nas últimas 3 semanas (ex. gripe, pneumonia, constipação)** *

Marcar apenas uma oval.

- ( ) Sim
- ( ) Não

14. **Qual a sua situação hormonal atual (apenas para o sexo feminino)**

Pode selecionar uma ou mais opções que melhor representem a sua situação atual
Marcar tudo o que for aplicável.

- [ ] Fase Reprodutiva
- [ ] Gravidez
- [ ] Menopausa
- [ ] Fez Histerectomia (remoção do útero)
- [ ] Fez ovariectomia (remoção do(s) ovário(s))
- [ ] Faz Terapia Hormonal de Substituição / Terapia de Compensação
- [ ] Outra: ______________________

15. **Toma algum tipo de medicação regularmente** *

(com ou sem receita médica)
Marcar apenas uma oval.

- Não *Passe para a pergunta 17.*
- Sim *Passe para a pergunta 16.*

Medicação

16 **Especifique o tipo de medicação que toma** *

Caso não saiba o nome do(s) medicamento(s), indique qual a sua finalidade (ex. hipertensão, depressão, diabetes,...)

Passe para a pergunta 17.

História Vocal

17. **Já teve alguma doença, acidente ou complicação cirúrgica que tenha afetado a sua voz?** *

Em caso afirmativo especifique na opção "outra"
Marcar apenas uma oval.

- Não
- Outra:

18. **Sente que a sua voz hoje está diferente do normal** *

Marcar apenas uma oval.

- Não
- Sim

19. **Caso considere que a sua voz está diferente, especifique:**

20. **Profissionalmente tem de falar (apenas para pessoas em idade ativa)**
Marcar apenas uma oval.

- Nunca
- Raramente
- Algumas vezes
- Muitas vezes
- Sempre

21. **Em que circunstâncias tem de falar profissionalmente (apenas para pessoas em idade ativa)**
Marcar tudo o que for aplicável.

- [ ] Durante muitas horas
- [ ] Sem pausas
- [ ] Em condições adversas (ex. ruído, ambiente poluído, variações de temperatura, etc)
- [ ] Não aplicável

22 **Canta** *
Marcar apenas uma oval.

- Nunca
- Raramente
- Algumas vezes
- Muitas vezes
- Sempre

23. **Tem algum tipo de treino vocal** *
Marcar apenas uma oval.

- Não *Passe para a pergunta 25.*
- Aulas de canto *Passe para a pergunta 24.*
- Aulas de representação *Passe para a pergunta 24.*
- Outra: ______________________ *Passe para a pergunta 24.*

Treino Vocal

24. **Número de anos de treino vocal** *

____

Passe para a pergunta 25.

Terapia da Fala

25. **Alguma vez teve Terapia da Fala** *

Marcar apenas uma oval.

- Sim
- Não

26. **Se sim, qual o motivo**

Otorrinolaringologia

27. **Alguma vez foi acompanhado em Otorrinolaringologia** *

Marcar apenas uma oval.

- Sim
- Não

28. **Se sim, qual o motivo**

Dispositivos de apoio

29 **Consegue ouvir** *

Marcar apenas uma oval

- Sim
- Com alguma dificuldade
- Não consigo ouvir quase nada

30. **Utiliza aparelho auditivo** *

Marcar apenas uma oval.

- Sim
- Não

31. **Consegue ver** *

Marcar apenas uma oval.

- Sim
- Com alguma dificuldade
- Não consigo ver quase nada

32. **Usa óculos ou lentes de contacto***

Marcar apenas uma oval.

- Sim
- Não

33. **Usa algum tipo de: ***

Pode assinalar mais do que uma opção:
Marcar tudo o que for aplicável.

- Prótese dentária fixa (ex. Implante, Coroa Dentária, Ponte, Faceta ou Prótese Adesiva)
- Prótese dentária removível (ex. Prótese Acrílica, Prótese Esquelética)
- Aparelho ortodôntico fixo
- Aparelho ortodôntico removível
- Aparelho Dentário Lingual
- Expansor palatino
- Não
- Outra: ____________________

Condições envolventes e hábitos

34 **Em casa, no emprego ou nas atividades de lazer está, muitas vezes, exposto(a) a : ***

Marcar tudo o que for aplicável.

- Pó
- Fumo (ex. tabaco, lareiras)
- Ruído
- Frio/Calor em excesso
- Humidade
- Produtos tóxicos
- Aquecimento central/ ar condicionado
- Pelo de animais
- Diferenças de temperaturas
- Não

35. **Hábitos tabágicos ***

Marcar apenas uma oval.

- Fumador
- Não fumador
- Ex-fumador (há mais de 5 anos)
- Ex-fumador (há menos de 5 anos)

36. **Quantos copos de água bebe por dia (aproximadamente)***

Marcar apenas uma oval.

- Não bebo
- 1 a 2 copos
- 3 a 4 copos
- 5 a 6 copos (+/- 1l)
- 7 a 8 copos (+/- 1,5l)
- Mais de 8 copos

37. **Com que frequência bebe bebidas alcoólicas** *

Marcar apenas uma oval.

- Nunca
- Ocasionalmente
- 1 vez por dia
- 2 vezes por dia
- Mais de 2 vezes por dia

38. **Com que frequência bebe bebidas com cafeína** *

Marcar apenas uma oval.

- Nunca
- Ocasionalmente
- 1 vez por dia
- 2 vezes por dia
- Mais de 2 vezes por dia

Fim

Muito obrigada pela sua colaboração!

HOSPITAL ANXIETY AND DEPRESSION SCALE (HADS)

This page intentionally contains only this sentence.

Escala de Ansiedade e Depressão Hospitalar

Pais-Ribeiro et al., 2007, Portuguese version of the HADS (The Hospital Anxiety and Depression Scale)

Este questionário foi construído para ajudar a saber como se sente. Pedimos-lhe que leia cada uma das perguntas e assinale o espaço anterior à resposta que melhor descreve a forma como se tem sentido na última semana.

Não demore muito tempo a pensar nas respostas. A sua reacção imediata a cada questão será provavelmente mais correcta do que uma resposta muito ponderada.

*Obrigatório

1. **Sinto-me tenso/a ou nervoso/a:** *

   Marcar apenas uma oval.

   - Quase sempre
   - Muitas vezes
   - Por vezes
   - Nunca

2. **Ainda sinto prazer nas coisas de que costumava gostar:** *

   Marcar apenas uma oval.

   - Tanto como antes
   - Não tanto agora
   - Só um pouco
   - Quase nada

3. **Tenho uma sensação de medo, como se algo terrível estivesse para acontecer:** *

   Marcar apenas uma oval.

   - Sim e muito forte
   - Sim, mas não muito forte
   - Um pouco, mas não me aflige
   - De modo algum

4. **Sou capaz de rir e ver o lado divertido das coisas:** *

   Marcar apenas uma oval.

   - Tanto como antes
   - Não tanto como antes
   - Muito menos agora
   - Nunca

5. **Tenho a cabeça cheia de preocupações:** *
Marcar apenas uma oval.

- A maior parte do tempo
- Muitas vezes
- Por vezes
- Quase nunca

6. **Sinto-me animado/a:** *
Marcar apenas uma oval.

- Nunca
- Poucas vezes
- De vez em quando
- Quase sempre

7. **Sou capaz de estar descontraidamente sentado/a e sentir-me relaxado/a:** *
Marcar apenas uma oval.

- Quase sempre
- Muitas vezes
- Por vezes
- Nunca

8. **Sinto-me mais lento/a, como se fizesse as coisas mais devagar:** *
Marcar apenas uma oval.

- Quase sempre
- Muitas vezes
- Por vezes
- Nunca

9. **Fico de tal forma apreensivo/a (com medo), que até sinto um aperto no estômago:** *
Marcar apenas uma oval.

- Nunca
- Por vezes
- Muitas vezes
- Quase sempre

10. **Perdi o interesse em cuidar do meu aspecto físico:** *
Marcar apenas uma oval.

- Completamente
- Não dou a atenção que devia
- Talvez cuide menos que antes
- Tenho o mesmo interesse de sempre

11. **Sinto-me de tal forma inquieto/a que não consigo estar parado/a:** *
Marcar apenas uma oval.

- Muito
- Bastante
- Não muito
- Nada

12. **Penso com prazer nas coisas que podem acontecer no futuro:** *
Marcar apenas uma oval.

- Tanto como antes
- Não tanto como antes
- Bastante menos agora
- Quase nunca

13. **De repente, tenho sensações de pânico:** *
Marcar apenas uma oval.

- Muitas vezes
- Bastantes vezes
- Por vezes
- Nunca

14. **Sou capaz de apreciar um bom livro ou um programa de rádio ou televisão:** *
Marcar apenas uma oval.

- Muitas vezes
- De vez em quando
- Poucas vezes
- Quase nunca

15. **Código do Participante** *
Fornecido pela investigadora

This page intentionally contains only this sentence.

ETHICAL APPROVAL (US DATABASE)

This page intentionally contains only this sentence.

COMISSÃO DE ÉTICA

da **Unidade Investigação em Ciências da Saúde: Enfermagem** (UICISA: E)

da **Escola Superior de Enfermagem de Coimbra** (ESEnfC)

Parecer Nº 639/ 12-2019

Título do Projecto: Estudo do correlato acústico-articulatório das vogais do Português-Europeu ao longo da idade.

Identificação das Proponentes

Nome(s): Catarina Alexandra Monteiro de Oliveira; António Joaquim da Silva Teixeira; Luciana Patrícia Martins Nunes Pereira Albuquerque; Ana Rita dos Santos Valente; Fábio Daniel Rodrigues Barros; Samuel de Sousa Silva; Paula Maria Vaz Martins; Daniela Maria Pias de Figueiredo

Filiação Institucional: Instituto de Engenharia Eletrónica e Informática da Universidade de Aveiro (IEETA)

Investigador Responsável/Orientador: Professora Doutora Catarina Alexandra Monteiro de Oliveira

Relator: Ana Margarida Abrantes

Parecer

O presente estudo parte integrante do projeto "Envelhecimento vocal: estudos das alterações segmentais e variações de ritmo em Português Europeu (VOX SENES)" (referência POCI-01-0145- FEDER-030820 (PTDC/LLT-LIN/30820/2017), financiado pela Fundação para a Ciência e Tecnologia, possui como objetivo principal objetivo o estudo das alterações de fala resultantes do envelhecimento. Assim, como objetivos secundários este projeto visa investigar:

i) a relação entre o posicionamento da língua na produção das vogais do Português-Europeu em contexto de palavra e a variação das frequências dos formantes ao longo da idade;
ii) a relação entre o posicionamento da língua na produção dos ditongos do Português-Europeu e a variação das frequências dos formantes ao longo da idade;
iii) a relação entre o posicionamento da língua na produção das vogais do Português-Europeu em contexto de leitura/discurso e a variação das frequências dos formantes ao longo da idade;
iv) a relação entre os movimentos da língua na produção das vogais do Português-Europeu em contexto de sílabas repetidas (diadococinésia) e a velocidade de fala ao longo da idade;

Todos estes objetivos serão concretizados através de um estudo em que os contornos da língua, obtidos por ultrassonografia serão sincronizados com o sinal acústico obtido.

Este estudo é definido como um estudo de caráter quantitativo e transversal a decorrer entre 1 de dezembro de 2019 e 1 de dezembro de 2021 com data prevista de colheita de dados entre 03 de fevereiro de 2020 e 1 de fevereiro de 2021.

A amostra será constituída por pelo menos 20 informantes, com idade superior a 18 anos. Os dados serão recolhidos no Instituto de Engenharia Eletrónica e Informática da Universidade de Aveiro.

COMISSÃO DE ÉTICA

da **Unidade Investigação em Ciências da Saúde: Enfermagem** (UICISA: E)

da **Escola Superior de Enfermagem de Coimbra** (ESEnfC)

Os instrumentos de recolha de dados são apresentados assim como o consentimento informado sendo garantida a confidencialidade, a voluntariedade e a autonomia das participantes. Não estão previstos qualquer custo ou necessidade de compensação para os participantes.

Sendo assim, somos do parecer que para o projeto poder ser aprovado sem restrições de natureza ética.

O relator: Ana Abrantes

Data: 22/01/2020 O Presidente da Comissão de Ética: Maria Florence Botelho

UNIDADE DE INVESTIGAÇÃO EM CIÊNCIAS DA SAÚDE ENFERMAGEM

Escola Superior de Enfermagem de Coimbra

FCT Fundação para a Ciência e a Tecnologia

CONSENT FORM (US DATABASE)

This page intentionally contains only this sentence.

"Estudo do correlato acústico-articulatório das vogais do Português-Europeu ao longo da idade.

Folha de Informação aos Participantes

Informações sobre o estudo:
O "Estudo do correlato acústico-articulatório das vogais do Português-Europeu ao longo da idade" constitui uma das tarefas do projeto "Envelhecimento vocal: estudos das alterações segmentais e variações de ritmo em Português Europeu (VOX SENES)" (referência POCI-01-0145-FEDER-030820 (PTDC/LLT-LIN/30820/2017), financiado pela Fundação para a Ciência e Tecnologia, a decorrer na Universidade de Aveiro – Instituto de Engenharia Eletrónica e Informática da Universidade de Aveiro (IEETA). O presente estudo insere-se, igualmente, no âmbito do plano de Doutoramento em Gerontologia e Geriatria (Universidade de Aveiro) da doutoranda Luciana Albuquerque, sob a orientação científica da Professora Doutora Catarina Oliveira, do Professor Doutor António Teixeira e da Professora Doutora Daniela Figueiredo.

Equipa de Investigação:
Catarina Oliveira (ESSUA/IEETA); António Teixeira (DETI/IEETA), Luciana Albuquerque (IEETA/CINTESIS-UA), Ana Rita Valente (IEETA), Fábio Barros (IEETA), Samuel Silva (DETI/IEETA), Paula Martins (ESSUA/IEETA), Daniela Figueiredo (ESSUA/CINTESISUA).

Objetivo do estudo:
Este tem como principal objetivo o estudo das alterações de fala (a nível segmental, supra-segmental, acústico e articulatório) resultantes do envelhecimento. Nomeadamente, investigar a relação entre o posicionamento da língua na produção das vogais e ditongos do Português-Europeu e a variação de parâmetros acústicos (ex. frequência fundamental, formantes e duração) ao longo da idade, através de um estudo em que os contornos da língua, obtidos por ultrassonografia, estão sincronizados com o sinal acústico.

População-Alvo:
O estudo destina-se a adultos saudáveis com idade superior a 18 anos.

Procedimento Específico:
Este estudo é constituído por duas etapas. Na primeira etapa, ser-lhe-á solicitado que preencha um questionário sociodemográfico em papel. Este tem como objetivo: 1) recolher dados sociodemográficos; 2) identificar fatores que influenciam a qualidade vocal.
Na segunda etapa, haverá uma recolha de dados acústicos e de ultrassonografia (obtenção dos contornos da língua). Inicialmente, ser-lhe-á colocado um estabilizador na cabeça, para permitir a estabilidade do transdutor de ultrassom (sonda). O transdutor de ultrassom ser-lhe-á colocado sob o queixo com o objetivo de captar imagens da superfície da língua, enquanto produz vogais sustentadas, palavras ou sílabas. Ser-lhe-á pedido que produza, num tom e ritmo de fala que considere normal, três vogais sustentadas e diferentes conjuntos de palavras, que serão repetidos 3 vezes. As palavras serão apresentadas num ecrã de computador com recurso a representação ortográfica. A recolha dos dados de fala será realizada utilizando um microfone conectado a uma placa de som externa. Poderá solicitar, a qualquer momento da gravação, a realização de um período de descanso.
Ser-lhe-á atribuído um código numérico, que permitirá associar os dados do questionário aos dados obtidos na recolha de dados biométricos.

Duração:
O presente estudo terá uma duração aproximada de 30 minutos (5 minutos para preenchimento do questionário, 5 minutos para colocação do estabilizador de cabeça e 20 minutos de gravação).

Natureza Voluntária da sua participação:
A sua participação é voluntária, pelo que poderá optar, a qualquer momento, por desistir do estudo, sem que daí advenham quaisquer prejuízos ou consequências.

Tem direito ao acesso, retificação e oposição ao tratamento dos seus dados. Se pretender exercer qualquer um destes direitos após ter submetido as suas respostas, por favor contate a investigadora responsável (Catarina Oliveira (coliveira@ua.pt)).

Riscos Associados:
Com base em estudos anteriores com procedimentos semelhantes, prevemos que a participação neste estudo não acarrete qualquer risco para o seu bem-estar físico e psicológico. Este estudo foi alvo de análise e aprovação pela Comissão de Ética da Escola Superior de Enfermagem de Coimbra.

Benefícios Associados:
Com a participação neste estudo estará a contribuir para aprofundar o conhecimento na área da linguística e do envelhecimento vocal. Este estudo não envolverá contrapartidas financeiras.

Confidencialidade e Anonimização:
A informação fornecida ou quaisquer dados recolhidos ao longo deste estudo serão usados apenas para fins de investigação científica, estando salvaguardada a total confidencialidade das informações recolhidas. Os dados pessoais sensíveis relacionados com as questões de saúde serão guardados pseudoanonimizados numa base de dados protegida com palavra-passe. Os dados de voz serão eliminados, previsivelmente, 3 anos após o término do seu estudo. Posteriormente, o projeto irá manter os dados biométricos (ultrasons) e os dados pessoais em formato anónimo. Os dados pessoais recolhidos serão guardados num servidor seguro da Universidade de Aveiro.

Responsáveis pelo Tratamento, acesso e partilha dos dados pseudonimizados:
Os responsáveis pelo tratamento dos dados são os investigadores do projeto acima identificados. Os dados pessoais não serão comunicados a nenhuma entidade nem há possibilidade de serem transferidos para países terceiros. Todos os elementos da equipa de investigação do projeto têm acesso à base de dados pseudonimizada. A base de dados anonimizada também poderá ser partilhada com revistas internacionais e apresentada em congressos científicos e outras publicações.

Contactos/ Esclarecimentos:
Caso deseje obter informações adicionais sobre este estudo, poderá contactar a investigadora Catarina Oliveira (coliveira@ua.pt) Se detetar algum problema pode contactar o encarregado de proteção de dados da UA (epd@ua.pt) ou a Comissão Nacional de Proteção de Dados (www.cnpd.pt).
O estudo dá cumprimento ao estipulado no Regulamento Geral de Proteção de Dados (RGPD), garantindo a segurança e confidencialidade de todos os dados facultados pelos participantes, em todas as fases do processo. O estudo segue também as recomendações da Declaração de Helsínquia para a investigação científica.

Declaração de Consentimento Informado

Declaro que:
Tenho 18 anos ou mais, que tomei conhecimento do objetivo do estudo e do que tenho de fazer para participar no mesmo. Declaro também que tive oportunidade de ler na íntegra este consentimento informado, que o considero explícito e que concordo com o seu conteúdo.
Fui informado/a que tenho o direito de recusar participar ou desistir em qualquer momento do estudo, e que essa recusa ou desistência não terão consequências para mim.
Foi-me garantida a confidencialidade da minha participação neste estudo.

Assim declaro que:

☐ **aceito**

☐ **não aceito**

participar na presente investigação, conduzida em estrita obediência ao Regulamento Geral da Proteção de Dados e da sua Lei de Execução Nacional.

Participante

Data: _____/_____/_____ ____
Assinatura

BACKGROUND QUESTIONNAIRE (US DATABASE)

This page intentionally contains only this sentence.

Envelhecimento vocal: estudos das alterações segmentais e variações de ritmo em Português Europeu (VOX SENES)

O projeto "Envelhecimento vocal: estudos das alterações segmentais e variações de ritmo em Português Europeu (VOX SENES)" (referência POCI-01-0145-FEDER-030820 (PTDC/LLTLIN/30820/2017) é um projeto financiado pela Fundação para a Ciência e Tecnologia, que tem como principal objetivo o estudo das alterações de fala (a nível segmental, suprasegmental, acústico e articulatório) resultantes do envelhecimento.
A sua participação neste estudo envolverá a gravação de uma amostra de fala, tendo em vista a obtenção de parâmetros acústicos (ex. frequência fundamental, formantes e duração das vogais). Simultaneamente, serão obtidas imagens dos articulatórios da língua através de ultrassonografia. Ser-lhe-á ainda solicitado que preencha o questionário de caracterização sociodemográfica que se segue.
Será garantida a confidencialidade e anonimato de todas as informações recolhidas, sendo apenas alvo de processamento os dados acústicos e articulatórios, bem como os dados referentes ao género e idade. De modo a garantir o anonimato, será utilizado um código numérico para identificação de cada um dos participantes, atribuído no momento de recolha da informação pessoal.
A sua participação é voluntária, pelo que poderá optar, a qualquer momento, por desistir do estudo, sem que daí advenham quaisquer prejuízos ou consequências.
A equipa de investigadores está ao seu dispor para responder a qualquer dúvida que considere pertinente.

Muito obrigada pela sua disponibilidade!

*Obrigatório

1. Código do Participante *

   Fornecido pela investigadora

   ____

2. Data de Nascimento *

   ____

   Exemplo: 7 de janeiro de 2019

3. Sexo *

Marcar apenas uma oval.

- Feminino
- Masculino

4. Habilitações académicas *

Marcar apenas uma oval.

- Não frequentou a escola ou não concluiu o 1.º Ciclo
- 1.º Ciclo (4.º ano de escolaridade)
- 2.º Ciclo (6.º ano de escolaridade)
- 3.º Ciclo ou formação equivalente (9.º ano de escolaridade)
- Ensino secundário ou formação equivalente (12.º ano de escolaridade)
- Ensino superior
- Outra:

5. Nasceu em que concelho *

6. Vive em que concelho *

7. Indique todos os locais (concelhos/ cidades/ países) em que tenha vivido durante pelo menos 2 anos. *

Indique também o número de anos que viveu em cada um dos locais que referir.

Informação sobre saúde

8. Como classifica o seu estado de saúde geral? *

Marcar apenas uma oval.

- Mau
- Médio
- Bom
- Excelente

9. Tem ou já teve algum (a): *

Pode assinalar mais do que uma opção:

Marcar tudo o que for aplicável.

- [ ] Traumatismo Cranioencefálico
- [ ] Acidente Vascular Cerebral (AVC)
- [ ] Doença Neurológica (ex. Esclerose Múltipla, Epilepsia, Parkinson)
- [ ] Perturbação Neurocognitiva (ex. Demência)
- [ ] Alteração Emocional (ex. depressão, fadiga crónica, estados de ansiedade, insónias)
- [ ] Cancro/Tumor de cabeça ou pescoço (ex. tiróide, garganta)
- [ ] Problema hormonal/ endócrino (ex. hipertiroidismo, hipotiroidismo)
- [ ] Azia crónica ou refluxo
- [ ] Perturbação de Fala/ Voz ou Linguagem (ex. gaguez, laringite crónica, paralisia das pregas vocais, afasia)
- [ ] Problema Respiratório crónico (ex. asma, rinite, sinusite)
- [ ] Cirurgia oral ou facial
- [ ] Cirurgia de cabeça ou pescoço
- [ ] Cirurgia do tórax
- [ ] Não

10. Sente alguma dificuldade a comer ou beber algum tipo de alimento? / Sente que se engasga com frequência? *

Em caso afirmativo especifique a frequência e as circunstâncias em que ocorre na opção "outra"

Marcar apenas uma oval.

- ( ) Não
- ( ) Outra: ____________________

11. Qual o seu peso atual (kg)? *

____

12. Qual a sua altura (cm)? *

____

13. Teve alguma infeção respiratória nas últimas 3 semanas (ex. gripe, pneumonia, constipação) *

Marcar apenas uma oval.

- ( ) Sim
- ( ) Não

14. Qual a sua situação hormonal atual (apenas para o sexo feminino)

Pode selecionar uma ou mais opções que melhor representem a sua situação atual

Marcar tudo o que for aplicável.

- [ ] Fase Reprodutiva
- [ ] Gravidez
- [ ] Menopausa
- [ ] Fez Histerectomia (remoção do útero)
- [ ] Fez ovariectomia (remoção do(s) ovário(s))
- [ ] Faz Terapia Hormonal de Substituição / Terapia de Compensação

Outra: [ ] ____________________

História Vocal

15. Já teve alguma doença, acidente ou complicação cirúrgica que tenha afetado a sua voz? *
Em caso afirmativo especifique na opção "outra"

Marcar apenas uma oval.

- Não
- Outra: ______________

16. Sente que a sua voz hoje está diferente do normal *

Marcar apenas uma oval.

- Não
- Sim

17. Caso considere que a sua voz está diferente, especifique:

____

18. Hábitos tabágicos *

Marcar apenas uma oval.

- Fumador
- Não fumador
- Ex-fumador (há mais de 5 anos)
- Ex-fumador (há menos de 5 anos)

Terapia da Fala

19. Alguma vez teve Terapia da Fala *

Marcar apenas uma oval.

- Sim
- Não

20. Se sim, qual o motivo

Otorrinolaringologia

21. Alguma vez foi acompanhado em Otorrinolaringologia *

Marcar apenas uma oval.

- Sim
- Não

22. Se sim, qual o motivo

Dispositivos de apoio

23. Consegue ouvir *

Marcar apenas uma oval.

- Sim
- Com alguma dificuldade
- Não consigo ouvir quase nada

24. Utiliza aparelho auditivo *

Marcar apenas uma oval.

- Sim
- Não

25. Consegue ver *

Marcar apenas uma oval.

- Sim
- Com alguma dificuldade
- Não consigo ver quase nada

26. Usa óculos ou lentes de contacto *

Marcar apenas uma oval.

- Sim
- Não

27. Usa algum tipo de: *

Pode assinalar mais do que uma opção:

Marcar tudo o que for aplicável.

- [ ] Prótese dentária fixa (ex. Implante, Coroa Dentária, Ponte, Faceta ou Prótese Adesiva)
- [ ] Prótese dentária removível (ex. Prótese Acrílica, Prótese Esquelética)
- [ ] Aparelho ortodôntico fixo
- [ ] Aparelho ortodôntico removível
- [ ] Aparelho Dentário Lingual
- [ ] Expansor palatino
- [ ] Não

Outra: [ ] ____________________

28. Caracterize a sua dentição (Indique se tem algum dente em falta e a sua localização): *

Muito obrigada pela sua colaboração!

COVID-19 CONTINGENCY PLAN

This page intentionally contains only this sentence.

Plano de prevenção e atuação na aquisição de dados de ultrassons no âmbito do "Estudo do correlato acústico-articulatório das vogais do Português-Europeu ao longo da idade."

(seguindo as Orientações nº 022/2020 de 01/05/2020 atualizada a 20/07/2020 e nº 006/2020 de 26/02/2020 da Direção Geral da Saúde (DGS))

1- Introdução

O presente documento dá a divulgar os pontos essenciais do Plano de Contingência relativo à recolha de dados que ocorrerá nas instalações do IEETA (Universidade de Aveiro), no âmbito da tarefa " Estudo do correlato acústico-articulatório das vogais do Português-Europeu ao longo da idade", para a Doença por Coronavírus (COVID-19) estabelecido pelos responsáveis do referido projeto, Professora Catarina Oliveira e Professor António Teixeira e pelos investigadores Luciana Albuquerque, Ana Rita Valente e Fábio Barros. Este plano fornece informação aos participantes sobre as medidas de prevenção e controlo desta infeção, e sobre os procedimentos e medidas a adotar perante a identificação de casos suspeitos e/ou confirmados.

O Plano de Contingência para a Doença por Coronavírus (COVID-19) foi desenvolvido com base nas orientações da Direção-Geral da Saúde (DGS) e na melhor evidência científica disponível até ao momento.

Os participantes na tarefa "Estudo do correlato acústico-articulatório das vogais do Português-Europeu ao longo da idade" serão informados sobre a doença por coronavírus (COVID19) e sobre as formas de evitar a transmissão, através dos meios mais adequados disponibilizados pela Universidade de Aveiro: Boletim Informativo, por correio eletrónico, afixação de cartazes nos espaços comuns, etc,. De igual modo, a informação sobre as recomendações e procedimentos estabelecidos no Plano de Contingência da Universidade de Aveiro para a Doença por Coronavírus (COVID-19) será amplamente divulgada, através dos meios mais adequados. A Universidade de Aveiro está comprometida com a proteção da saúde e a segurança dos seus colaboradores, tendo também um papel importante a desempenhar na limitação do impacto negativo deste surto na comunidade, face às valências de conhecimento que detêm em diversas áreas.

> *"O plano de Prevenção e Atuação Face à COVID-19 da Universidade de Aveiro (UA) é uma ferramenta estratégica que define as medidas de prevenção e mitigação dos riscos associados à propagação do vírus SARS-CoV-2.*
>
> *O presente plano segue as orientações da Organização Mundial da Saúde (OMS) e do Centro Europeu de Prevenção e Controlo de Doenças (CEPCD), bem como o Plano de Contingência Nacional e as Orientações emanadas pela Direção Geral de Saúde (DGS).*
>
> *A atual situação relativa à doença COVID-19 foi considerada pelo CEPCD como de impacto elevado e provável propagação global, tendo levado a OMS a declarar situação de Pandemia. Nesse sentido, foram decretadas medidas de contingência e divulgadas práticas para mitigação da propagação do referido vírus, com vista à minimização dos riscos de contágio e propagação da doença COVID-19." (Plano de contingência da UA)*

2- Objetivos do presente Plano de prevenção e atuação

▪ Minimizar o impacto da COVID-19 nos participantes e nos investigadores envolvidos no "Estudo do correlato acústico-articulatório das vogais do Português-Europeu ao longo da idade";

▪ Colocar em prática medidas de contenção adequadas face ao nível de risco;

▪ Assegurar a atempada monitorização da situação, bem como a recolha e comunicação de informação relativa ao COVID-19;

▪ Assegurar o decorrer do estudo de acordo e em função do nível de risco.

3- A doença por coronavírus (COVID-19) e a sua transmissão

A COVID-19 foi considerada uma Pandemia a 11 de março de 2020 pela Organização Mundial de Saúde. Em Portugal, as medidas de Saúde Pública têm sido implementadas de acordo com as várias fases de preparação e resposta a situações epidémicas, por forma a diminuir progressivamente a transmissão do vírus, prestar os cuidados de saúde adequados a todos os doentes e proteger a Saúde Pública.

A COVID-19 é uma doença causada pela infeção pelo novo Coronavírus (SARS-CoV-2). Os sinais e sintomas da COVID-19 variam em gravidade, desde a ausência de sintomas (sendo assintomáticos) até febre (temperatura ≥ 38.0ºC), tosse, dificuldade respiratória, odinofagia (dor de garganta), dores musculares generalizadas, cefaleias (dores de cabeça), fraqueza, e, com menor frequência, náuseas/vómitos e diarreia. Nos casos mais graves pode-se verificar pneumonia grave, síndrome respiratória aguda grave, septicémia, choque sético e eventual morte. Os dados mostram que o agravamento da situação clínica pode ocorrer rapidamente, geralmente durante a segunda semana da doença. Recentemente, foi também verificada anosmia (perda do olfato) e em alguns casos a perda do paladar, como sintomas da COVID-19.

Com base na evidência científica atual, este vírus transmite-se principalmente através de:

− Contacto direto: disseminação de gotículas respiratórias, produzidas quando uma pessoa infetada tosse, espirra ou fala, que podem ser inaladas ou pousar na boca, nariz ou olhos de pessoas que estão próximas (< 2 metros).

− Contacto indireto: contacto das mãos com uma superfície ou objeto contaminado com SARS-CoV-2 e, em seguida, com a boca, nariz ou olhos.

4- Procedimentos Gerais

A recolha de dados à qual diz respeito o presente plano de prevenção acarreta um conjunto de exigências (i.e., proximidade dos colaboradores com o participante, não uso de máscara no momento de gravação), existindo gotículas respiratórias e aerossóis que podem ser criados durante os procedimentos. Por isso, devem ser tomadas medidas adicionais para assegurar uma minimização da transmissão do vírus.

Esta orientação tem em conta a fase de transmissão comunitária da Infeção por SARS-CoV-2 e poderá ser revista a qualquer momento, em função da evolução do conhecimento científico e da situação epidemiológica.

Colocando em prática as orientações da DGS, não será realizada nenhuma recolha de dados de ultrassons sem um prévio contacto por via remota (telemóvel, e-mail, ou outro meio que permita a comunicação com o participante). Assim como todos os investigadores terão a formação/informação necessária para agir de acordo com o plano de contingência.

5- Triagem Prévia à recolha de dados

1. Antes da realização da recolha de dados será feita uma triagem prévia para que o participante seja avaliado quanto à presença de sintomas sugestivos de COVID-19:

a. Questionar o participante relativamente à presença de quadro respiratório agudo com tosse (de novo ou agravamento da tosse habitual), ou febre (temperatura ≥ 38.0ºC), ou dispneia/dificuldade respiratória nos últimos 14 dias.

b. Questionar o participante se esteve em contacto com um caso suspeito ou confirmado de COVID-19. Em caso afirmativo, questionar se ainda se encontra em período de isolamento (14 dias).

c. Questionar o participante se teve diagnóstico prévio de COVID-19. Em caso afirmativo, questionar se ainda se encontra em período de isolamento (14 dias).

d. Questionar o participante se viajou para algum país de risco ou zona com transmissão ativa da COVID-19 nos últimos 14 dias.

2. Se o participante referir sintomas sugestivos de COVID-19, será contactada a Linha SNS24 (808 24 24 24), nos termos da Norma 004/2020 da DGS.

3. Nestes casos não irá ocorrer a recolha de dados de ultrassons, sendo adiada a participação no estudo e, remarcada nova data de recolha de dados para depois da recuperação/cura do participante, respeitando o período de isolamento.

6- Definição da área de isolamento

A colocação de um colaborador ou visitante suspeito de infeção por COVID-19 numa área de isolamento visa impedir que outros colaboradores possam ser expostos e infetados. Esta medida tem como principal objetivo evitar a propagação de uma doença transmissível.

Na Universidade de Aveiro foi definida a seguinte área de isolamento: Campus Universitário de Santiago: Sala 6.1.55 (Antigo gabinete médico na Zona Técnica Central – "catacumbas") (área de isolamento mais próxima ao IEETA). Esta encontra-se equipada com: telefone, cadeira ou marquesa, água, solução antisséptica de base alcoólica, máscaras cirúrgicas e luvas descartáveis, toalhetes de papel, contentor de resíduos, termómetro e equipamentos de limpeza.

Os responsáveis do estudo (Luciana Albuquerque, Ana Rita Valente, Fábio Barros), informarão o Grupo de Acompanhamento da COVID-19 da UA sobre a existência de um caso suspeito (Contacto: covid19@ua.pt), de modo a proceder ao seu encaminhamento para a área de isolamento especificada.

7- Procedimentos a seguir <u>antes</u> da recolha de dados

1. Agendar/iniciar a sessão de recolha de dados de ultrassons apenas após responder à triagem prévia.

2. Explicar ao participante, quando for marcada a sessão de recolha de dados, os procedimentos de controlo e prevenção de infeção que estão implementados, nomeadamente:

a. Desaconselhar a presença de acompanhante;

i. Se necessária a presença do acompanhante, este deve, preferencialmente, aguardar fora das instalações, ou então permanecer na sala de espera/corredor (com máscara cirúrgica colocada);

b. Se o participante se apresentar nas instalações do IEETA sem máscara facial, será colocada uma à entrada, fornecida no local;

c. Explicar ao participante a necessidade de higienização das mãos com solução alcoólica disponível no local, à chegada e à saída;

d. Explicar ao participante a necessidade de distanciamento de, pelo menos, 2 metros das outras pessoas que possam estar presentes;

e. Explicar ao participante a necessidade de evitar tocar em superfícies desnecessariamente;

f. Explicar ao utente a necessidade de evitar o uso de adereços, como anéis, pulseiras, colares, brincos e relógios.

3. Assegurar, sempre que possível, uma boa ventilação da sala, preferencialmente com ventilação natural, através da abertura de portas ou janelas.

4. Desinfetar as superfícies, dando especial atenção às de toque frequente, de acordo com a Orientação 014/2020 da DGS (maçanetas de portas, interruptores de luz, teclado de computador e equipamento de ultrassons, mesas, cadeiras, entre outros).

5. Assegurar a correta higienização do espaço recorrendo à utilização de película aderente em equipamentos de difícil limpeza (ex. teclado de computador).

6. Assegurar a correta desinfeção das espátulas. No caso de espátulas descartáveis, estas serão depositados sobre um tabuleiro descartável e pulverizados com álcool a 70º ou submersas numa solução de cloro (20 mL de solução de hipoclorito de sódio a 5% em 980mL de água) antes de serem embaladas individualmente. No caso das espátulas reutilizáveis, estas passarão pelo processo mencionado anteriormente, antes da esterilização em autoclave.

7. Preparar antecipadamente todo o material que seja necessário utilizar durante a sessão de recolha de dados, para evitar a circulação de pessoas e a abertura de gavetas. Sendo a lista de material necessário a seguinte:

a. Soluções desinfetantes

b. Papel absorvente

c. Gel ecográfico

d. Capacete de aquisição

e. Espátula para traçado de plano oclusal

f. Luvas

g. Informação ao participante, consentimento informado, questionário e caneta

h. Cobertura de espuma do microfone desinfetada.

8. Retirar todos os adereços, como anéis, pulseiras, colares, brincos e relógios para gravar os participantes do estudo.

9. Os investigadores do estudo presentes na sala devem ter todo o equipamento de proteção individual (EPI) colocado, antes de entrar na sala.

8- Procedimentos durante a recolha de dados

1. Garantir que os objetos de uso pessoal não estão expostos durante a recolha de dados e que são alvo de uma desinfeção regular.

2. Será apresentada uma bandeja ao participante, previamente higienizada, na qual estará uma espátula, individualmente embalada, que deve ser colocada na boca pelo próprio participante quando instruído a tal pelo investigador.

3. O participante deve ser instruído a manter a máscara facial colocada sempre que indicado.

4. Os investigadores devem permanecer, sempre que possível, afastados do participante em estudo durante a produção do corpus de fala.

5. Solicitar ao participante que higienize as mãos antes e no fim da recolha de dados.

9- Procedimentos após a recolha de dados

1. Limpar e desinfetar após a recolha de dados de cada participante todas as superfícies e o ambiente de trabalho, utilizando um desinfetante multissuperfícies perfumado (X-DEF P), adequado para a desinfeção geral de superfícies de todo o tipo (paredes, sanitários, mesas de trabalho, móveis, caixotes de lixo, corrimões, maçanetas de porta, etc), e toalhetes humedecidos em desinfetante ou em álcool para a limpeza rápida das superfícies de toque frequente.

2. Remover a cobertura de espuma do microfone e colocar para desinfeção.

3. Fazer a renovação do ar do gabinete no final da recolha de dados de cada participante.

4. Manter a sala desocupada durante 1 hora entre participantes.

5. Ter precauções redobradas no manuseamento dos equipamentos (aparelho de ultrassonografia portátil (Mindray DP6900), sonda endocavitária (65EC10EA), capacete de estabilização da sonda, microfone, computador, monitores, teclados e espátulas), assegurando a sua efetiva desinfeção.

6. Seguir escrupulosamente todos os procedimentos universais de esterilização e desinfeção.

7. Na realização dos procedimentos de desinfeção da sala, após a recolha de dados, o investigador deve retirar o primeiro par de luvas e só retirar o restante EPI, após o acondicionamento de todo o material utilizado.

8. Deitar todos os EPI e material descartável nos contentores de lixo destinados para esse fim.

10- Equipamento de Proteção Individual (EPI)

1. O investigador adstrito exclusivamente à receção deverá estar equipado com máscara FFP2 (N95), viseira, avental descartável sobre a farda de uso clínico e luvas.

2. O investigador adstrito exclusivamente ao controlo do software de recolha de dados deverá estar equipado com máscara FFP2 (N95), viseira, avental descartável sobre a roupa/farda de uso clínico e luvas.

4. Se reutilizáveis, as viseiras devem ser desinfetadas antes e depois de cada recolha de dados.

5. Ao remover o EPI, deve ser retirado o 1º par de luvas antes remover o restante EPI, e deixar o 2º par de luvas para o final.

11- Material Reutilizável

1. As viseiras, devem ser pulverizadas com álcool a 70º sobre um tabuleiro, nos termos da Orientação 014/2020 da DGS.

2. As peças de roupa do EPI laváveis devem ser retiradas sem sacudir, enroladas no sentido de dentro para fora, e acondicionadas em saco impermeável, fechando-o bem até ao local de lavagem e deposite a roupa diretamente para dentro da máquina.

3. As roupas devem ser lavadas à temperatura mais alta que puderem suportar (pelo menos a 60ºC durante 30 minutos), ou se a roupa não puder ser lavada a quente, deve ser lavada na máquina a temperatura entre 30-40ºC, com um desinfetante apropriado a este tipo de roupa e compatibilidade com a máquina.

12- Protocolo de recolha de dados de ultrassom com as adaptações relacionadas com o plano de contingência

1. O investigador responsável por controlar o software de gravação deve ocupar o seu lugar antes da entrada do participante na sala;

2. O investigador responsável por preparar o participante para a recolha de dados deve:

a. Posicionar o participante;

b. Facultar-lhe todas as informações necessárias sobre o estudo;

c. Pedir ao participante para assinar o consentimento informado;

d. Colocar o estabilizador de cabeça ao participante;

e. Posicionar/ reposicionar a sonda no participante sempre que necessário;

f. Manter-se no exterior da sala durante a gravação do corpus de fala.

3. Dar início à recolha do corpus de fala.