A Longitudinal Study of Speech Acoustics in Older French Females: Analysis of the Filler Particle 'euh' across Utterance Positions

Fuchs, Susanne; Koenig, Laura L.; Gerstenberg, Annette

Veröffentlichungsversion / Published Version
Zeitschriftenartikel / journal article

Article

A Longitudinal Study of Speech Acoustics in Older French Females: Analysis of the Filler Particle *euh* across Utterance Positions

Susanne Fuchs [1],*[iD], Laura L. Koenig [2] and Annette Gerstenberg [3][iD]

1   Laboratory Phonology Group, Leibniz Centre General Linguistics, 10117 Berlin, Germany
2   Communication Sciences and Disorders, Adelphi University, Garden City, NY 11530, USA; koenig@haskins.yale.edu
3   Department of Romance Languages, University Potsdam, 14469 Potsdam, Germany; gerstenberg@uni-potsdam.de
*   Correspondence: fuchs@leibniz-zas.de

Abstract: Aging in speech production is a multidimensional process. Biological, cognitive, social, and communicative factors can change over time, stay relatively stable, or may even compensate for each other. In this longitudinal work, we focus on stability and change at the laryngeal and supralaryngeal levels in the discourse particle *euh* produced by 10 older French-speaking females at two times, 10 years apart. Recognizing the multiple discourse roles of *euh*, we divided out occurrences according to utterance position. We quantified the frequency of *euh*, and evaluated acoustic changes in formants, fundamental frequency, and voice quality across time and utterance position. Results showed that *euh* frequency was stable with age. The only acoustic measure that revealed an age effect was harmonics-to-noise ratio, showing less noise at older ages. Other measures mostly varied with utterance position, sometimes in interaction with age. Some voice quality changes could reflect laryngeal adjustments that provide for airflow conservation utterance-finally. The data suggest that aging effects may be evident in some prosodic positions (e.g., utterance-final position), but not others (utterance-initial position). Thus, it is essential to consider the interactions among these factors in future work and not assume that vocal aging is evident throughout the signal.

Keywords: aging; prosody; voice quality; fundamental frequency; formants; filler particles

1. Introduction

Human aging is a multidimensional process that impacts anatomy, physiology, linguistic properties, communication, and cognition. Work on vocal aging suggests that the timing of changes may vary across individuals (Goozée et al. 1998). For female speakers, authors have given considerable attention to the effects of menopause (see summary in Lenell et al. 2019). This raises the question of whether we can still trace vocal aging in older females (beyond their 60's) or if patterns are relatively stable. In connected speech, a voice may also have different characteristics depending on the analysis unit, i.e., what portion of the speech stream we are analyzing or listening to. In combination, physical, sociolinguistic, and prosodic variables may make it hard to tease apart the factors contributing to age-related changes or compensating for them. As examples, Smorenburg and Heeren (2020) observed that speaker specific information varied with phonological and syllabic context and Weirich (2012) found that physiological aspects in speech production were visible in unstressed syllables, but not in stressed ones. These findings indicate that speaker information in speech sounds is not necessarily the same across linguistic contexts and prosodic positions.

To assess stability and changes in speech and language use, we employ here a longitudinal design to control for the numerous variables that can complicate cross-sectional

studies. Much of the work cited below was in fact cross-sectional (exceptions being Decoster and Debruyne 1997; Endres et al. 1971; Gerstenberg et al. 2018; Harrington et al. 2000; Keszler and Bóna 2019; Reubold et al. 2010; Russell et al. 1995). Cross-sectional designs have obvious practical advantages over longitudinal ones: They can be carried out in a relatively short time frame and do not suffer from issues of participant attrition. Comparing individuals across generations, as is necessary in a cross-sectional design, carries its own threats to validity, however. Along with simple random variation across participants, individual conditions of medical care, diet, and lifestyle factors such as smoking are not controlled, nor is the local impact of social and cultural changes that can affect linguistic and pragmatic variables. Thus, behavioral comparisons across generations may not necessarily lead to the same results as across the lifespans of individuals. We are focusing on a decade in the lifespan of the older generation (all of them aged 70+ at the second time point). As Fougeron et al. (2021) state, dynamics of vocal aging do not unfold evenly over the lifespan.

In addition, in contrast to most previous work, we focus on spontaneous speech, which yields greater ecological validity than read speech or sustained vowels. We specifically take acoustic measures from the filler particle *euh* in 10 older French females recorded 10 years apart. This particle has been widely investigated acoustically and occurs frequently in spontaneous speech (Candea 2000, p. 79), ensuring a high degree of inter-speaker comparability. While no empirical evidence was provided, Duez (2001) proposed that the acoustics of *euh* are "strongly linked to absolute physiological aspects of speech production" (p. 44). However, acoustic realizations may also depend on the position of *euh* in the utterance, i.e., prosody and contextual information may affect how the filler particle is realized (Shriberg 2001).

Our acoustic analysis covers formants (F1, F2) and fundamental frequency (f0), which have been widely studied, and voice quality, which has been assessed less frequently (but see the recent work by Fougeron et al. 2021; Karlsson and Hartelius 2021). All parameters can reflect physical effects of aging but may also be conditioned by prosody, specifically whether the filler occurs pre-pausally, post-pausally, or within a speech unit.

1.1. Filler Particles in Connected Speech

Filler particles are among the most frequent features in spontaneous speech, providing a convenient comparison across and within speakers. The literature on filler particles is quite broad, so we will first start with a definition of filler particles and then go into more details about their frequency, occurrence within discourse, and acoustic characteristics. Whenever possible, results from studies on aging will be introduced.

1.1.1. Definition

French *euh* can be categorized among non-lexical filler particles, defined by Belz (2021, p. 4). They have often been discussed in the context of pausing (Grosman et al. 2018). In this paper, we adopt 'filler particle' (FP) instead of other previously used terms such as hesitation or filled pause, which may carry negative connotations or fail to capture the multiple discourse functions of FPs. Ferreira and Bailey (2004) note that these particles (together with other phenomena) were almost ignored in linguistic research in the past but have received much more attention in the last decades. Filler particles have also been explored in various languages (see the comprehensive summary in Belz 2021, p. 14).

1.1.2. Filler Particles in Discourse: Frequency and Location

Euh is the most frequent filler in spoken French, and its discourse functions are multifaceted. Filler particles do not occur randomly. Rochet-Capellan and Fuchs (2013) did an analysis of fillers in German with respect to breath cycles. On average, 40% of all breathing cycles were realized with a filler particle, but results were highly speaker specific. If a filler occurred, again in 40% of cases, it was produced just after inhalation, at the beginning of speech. These results are congruent with previous work investigating

fillers with respect to respiration (Schönle and Conrad 1985). In French, the filler particle *euh* is often preceded or followed by a silent pause, both cases occurring with the same frequency. However, Grosman et al. (2018) observed that pauses preceding *euh* were shorter than pauses following *euh*, which could indicate that *euh* + pause represents an example of 'disfluency'. Gósy and Silber-Varod (2021) recently observed a comparable duration for young (but not older) speakers of Hungarian. More generally, the durational difference between these two contexts for *euh* could suggest varying discourse functions. *Euh* can have different roles according to its place in the information structure and context (e.g., following connectors such as *donc* 'so' or *mais* 'but'; Morel and Danon-Boileau 1998, pp. 82–83). Shriberg (2001) showed that the use of filler particles can be related to a variety of ecological factors, including communicative context, syntactic structure, individual behavior, and personal factors such as sex along with the age of the speaker.

Empirical findings for the frequency of fillers in spontaneous speech are diverse. Some studies on aging have reported higher frequency and longer durations of filled and unfilled pauses with age (e.g., Oyer and Deal 1985; Horton et al. 2010 for *uh* up to the age of 68). Bolly et al. (2016) studied *euh* as one of nine linguistic phenomena of disfluency in the Valibel corpus of French. For the entire series, they found a positive correlation between disfluency counts and speaker age.

No change in filler frequency was reported in a longitudinal study of Hungarian speakers. Keszler and Bóna (2019) analyzed the occurrence of the filler particle at three successive speaker ages, around 60, 70, and 75+ years. The authors noted individual differences but could not find an overall age effect. Moreover, the normalized number of fillers also did not display any difference in comparison to younger Hungarian speakers. Stability in the frequency of fillers has additionally been found in cross-sectional studies, such as Searl et al. (2002) who analyzed speaker groups at 70, 80, 90, and even 100+ years old.

Finally, some studies find fewer fillers with older age (Gósy and Silber-Varod 2021; Maxim et al. 1994, p. 112). In a seven-year comparison of LangAge speakers (the same corpus used here) the normalized frequency of *euh* significantly decreased with age (Gerstenberg 2015) while individual variations were also found. Taschenberger et al. (2019) studied the filler in different age populations interacting with an interlocutor and with no background noise, non-speech background noise or background speech. While on average no changes in the frequency of fillers were found, under noisy conditions, fewer filler particles were produced in older age. Gall (2019, p. 142) obtained similar results, with a 10% decrease of the filler particle in German in interviews with early retirees recorded 12–14 years apart.

The very notion FPs as an instance of disfluency is challenged by the results of Bortfeld et al. (2001) who found different distributions for 'disfluency' types: repeats and restarts were sensitive to utterance length, whereas filler particles varied with conversational roles. The authors proposed that some FPs could reflect or support coordination between conversation partners and not represent hesitations or repairs. Along similar lines, Horton et al. (2010, p. 711) found a positive correlation in the frequency of *uh* with age, but a negative correlation with *um*. They suggested that the increased use of *uh* may be related to word-finding problems, whereas the decreasing use of *um* could be related to changes in sentence-planning strategies. To sum up, contemporary work suggests that FPs function in diverse ways, and do not necessarily correspond to slowing, disfluency, or cognitive decline but are rather a general property of spontaneous speech.

1.1.3. Acoustic Characteristics of Filler Particles

Various studies have been carried out to investigate the acoustic properties of filler particles, considering language and speaker specificity. In a crosslinguistic study, Candea et al. (2005) reported language specificities for formants, but similar patterns regarding duration and f0. The results for formant frequencies are in line with the findings of Belz (2020). Reviewing the literature, he summarized that formant characteristics of the

filler particle correspond to /ə/ in English and /e/ in Spanish (see also Belz 2021, p. 17 for a comprehensive overview). His analysis showed that German fillers had formant values comparable to those of both /œ/ and /ɐ/ in lexical words, which may be similar to the French filler *euh*. Vowels in filler particles also had larger variability in formant frequencies than the vowels in lexical words that were recorded as references. Finally, Karpiński (2013) reported significantly different formant values across Polish speaking participants, i.e., considerable cross-talker variability, including the expected male-female differences.

According to our knowledge, less is known about the acoustic realizations of filler particles in the older speaker population. Gósy et al. (2014) analyzed Hungarian speakers in three age groups: 9-year-old children, young adults in their 20s, and older speakers between 75 and 90 years. Age group influenced F1 and F2 values: F1 values decreased with age but F2 values decreased from childhood to young adulthood, then increased again for the older adults.

Results for fundamental frequency in FPs differ across studies. As noted above, Candea et al. (2005) observed similar values of f0 across languages. Karpiński's (2013) data showed a limited f0 range. While Duez (2001) found speaker specific and relatively stable f0 values independent of the position in an utterance for French, Shriberg and Lickley (1993) found that f0 values for clause-internal fillers in English were dependent on the preceding f0 peak, i.e., they were sensitive to prosodic context. Morel and Danon-Boileau (1998, p. 82) noted that f0 of *euh* mainly corresponded to the "unmarked" level, i.e., it lay within the lower f0 range of the speaker. However, the comparability is limited, as varying observations for f0 might arise from typological differences of the analyzed languages, and from differences in the annotation scheme with the respective exclusion or inclusion criteria accounting for prosodic position. For example, some fillers occurring at the end of an utterance, just before a pause, may not be considered in f0 analyses, because they are realized with creaky voice or phonation may be completely absent (e.g., Belz 2021; Karpiński 2013).

1.2. Aging Effects on Speech Anatomy, Physiology, and Acoustics

The literature on vocal aging is likewise extensive and diverse in terms of speech tasks and age groups assessed. Until rather recently, it was heavily dominated by studies of English. To constrain the review, we will exclude studies that only included men and did not assess aging beyond age 50. Additionally, we did not systematically include studies of pathological aging; that is, the papers we cover focused on speakers without severe cognitive or physiological impairments.

We also note that we did not seek, in this study, to provide an exhaustive set of measures that might vary with age, but were particularly interested in laryngeal measures (viz., voice quality and f0). One frequently-discussed aspect of aging that we omit here is slowed speech or articulation rate. Readers are referred to the comprehensive studies of Schötz (2006) and Fougeron et al. (2021) for results on duration and other measures not included here.

1.2.1. Respiratory System

The respiratory system is subject to calcification of cartilages, an increase in stiffness, and a reduction in compliance with aging (Estenne et al. 1985; Segre 1971). On average, older persons may need to increase inspiratory and/or expiratory volumes compared to younger ones when speaking (Sperry and Klich 1992) to compensate for reduced vital capacities and increased residual volumes (Frank et al. 1957; Hoit and Hixon 1987). Older adults have been found to use a greater percentage of vital capacity per syllable and produce fewer syllables per breath group (Hoit and Hixon 1987; see also Gerstenberg et al. 2018). In a large-scale study employing the same corpus of spoken French used here, Gerstenberg (2011) observed that interpausal units (measured in seconds) decreased with age, especially in the female speakers.

1.2.2. Upper Vocal Tract

Anatomical data suggest that the supralaryngeal cavities may enlarge with aging. Skull dimensions increase into advanced age (Israel 1973; Lasker 1953), and the vocal tract may be lengthened by lowering the larynx (Segre 1971). Interestingly, Xue and Hao (2003) documented a chronological increase in length and volume of oral (but not pharyngeal) cavity length and volume for males and females, in contrast to earlier pilot data suggesting larger pharyngeal cavities in females as a function of age (Xue et al. 1999).

Increases in vocal tract size should lead to lowered formant frequencies, and several studies have reported a lowered F1 with advancing age (e.g., Albuquerque et al. 2020; Endres et al. 1971; Linville and Fisher 1985; Linville and Rens 2001; Reubold et al. 2010; Scukanec et al. 1991). One could expect the effects of larger vocal tract dimensions to extend to other formants as well, but the data here are more limited. Some reports exist for F2, but few studies report on F3, F4, and higher formant frequencies. The long-term average spectrum results of Linville and Rens (2001) showed lowered values of F1, F2, and F3, but the percentage decrease was largest for F1, and for men the results were only significant for F1. Linville and Rens attributed more extreme effects on F1 to a localized increase in posterior regions of the vocal tract and speculated that larger effects in females might reflect osteoporosis, more extreme thinning of inter-vertebral disks, and/or a greater susceptibility to weakened muscular support of the larynx. It should be noted, however, that formant changes with aging can reflect other factors. Reubold et al. (2010) interpreted F1 changes over time as an adaptation to f0 lowering with the purpose of maintaining f0–F1 relationships. Further, formant changes with age are not always consistent across vowels (see summary in Eichhorn et al. 2018). Some aspects of chronological change, especially across speakers, could reflect sound change (e.g., Reubold et al. 2010) and/or stylistic factors that may change with age and interact with gender. As one possible example, Fougeron et al. (2021) observed that measures of F1 and F2 ranges differed in older males and females: Males showed no significant changes in F1 range with age, whereas females did; conversely, F2 ranges changed with age in males but not females.

1.2.3. Larynx

Chronological aging effects on the larynx have been studied extensively. The thyroid cartilage and the articular surfaces of the arytenoids demonstrate calcification, ossification, and changes in the balance and organization of collagen, resulting in lower compliance (e.g., Hirano et al. 1983; Kahane 1987b, 1988; Segre 1971; Turk and Hogg 1993). The vocal folds themselves undergo a decrease in innervation and blood supply and changes in the quantity of muscular, collagenous, and elastic fibers (Hammond et al. 1998; Hammond et al. 2000; Michel et al. 1987; Segre 1971). Such changes may negatively impact control of vocal fold position and tension (Kahane 1987a; Paulsen and Tillmann 1998). Such effects are, on average, less extreme in females than males.

Laryngoscopic observations of the vocal folds in healthy older females have documented bowing, glottal gaps, vocal fold atrophy, edema, asymmetry, aperiodicity, stiffness, and reduced amplitude of vibration (Biever and Bless 1989; Honjo and Isshiki 1980; Lundy et al. 1998; Pontes et al. 2005). Some degree of incomplete glottal closure is common in females throughout adulthood (e.g., Linville 1992), but the position of the gap along the vocal folds may change (Biever and Bless 1989; Linville 1992; cf. also Yamauchi et al. 2014). In combination, these laryngeal changes could plausibly affect the fundamental frequency (f0), its stability, and measures of voice quality. Reduced laryngeal efficiency owing to vocal fold bowing may also lead speakers to adjust their utterances and/or laryngeal settings in light of aerodynamic requirements for speech.

Acoustic studies of females mostly report f0 lowering at advanced ages (Benjamin 1981; Brown et al. 1989; Dehquan et al. 2012; Ferrand 2002; Higgins and Saxman 1991; Honjo and Isshiki 1980; Reubold et al. 2010; Russell et al. 1995; Stoicheff 1981; Xue and Deliyski 2001). When one looks at females post-menopause, the results are somewhat more complex. Schötz (2006), assessing six words, found that f0 decreased until about age

50, followed by a slight increase to age 70, followed by a slight decrease. This database included 5 or 6 speakers per year for speakers in their 70's, but results for ages 80 and higher were based on 12 speakers. Stathopoulos et al. (2011) measured f0 in sustained /a/ for 6 speakers per decade up to age 50, after that 7–11 speakers per decade in their 60's, 70's, and 80's, and 4 speakers in their 90's. They similarly observed a decrease in f0 (to about age 60), followed by a slight increase. Finally, Fougeron et al. (2021) measured f0 in a sentence for 265 women, with fuller representation (37 or more speakers) per decade age 60 or older. They observed that f0 decreased to about age 40 and then remained stable. It is difficult to disentangle the effects of sampling and speech materials here. Still, it does appear that age-related f0 decreases for females reported in past work mainly reflect younger to middle-aged speakers, and that f0 in older females may be stable or show a slight increase.

Within-speaker f0 standard deviations have been found to increase with age (Brown et al. 1989; Linville and Fisher 1985; Stoicheff 1981; Xue and Deliyski 2001; cf. also Fougeron et al. 2021, who reported on the coefficient of variation of f0). For voice quality, most studies have focused on measures that correspond perceptually to a hoarse or harsh quality, viz. jitter and shimmer. Whereas shimmer seems to increase with age (Biever and Bless 1989; Dehquan et al. 2012; Fougeron et al. 2021; Xue and Deliyski 2001), results for jitter are conflicting: Some authors report higher aperiodicity in older females (Dehquan et al. 2012; Fougeron et al. 2021; Xue and Deliyski 2001) but many find no age effects (Biever and Bless 1989; Brown et al. 1989; Ferrand 2002; Linville and Fisher 1985; Schötz 2006). Other measures of aperiodicity include harmonics (or signal) to noise ratio (HNR) and cepstral peak prominence (CPP). Schultz et al. (2021) recently reported that the standard deviation of CPP contributed to the prediction of chronological age in males and females age 50–92 and noted that CPP might correlate with the use of creaky voice. Fougeron et al. (2021) observed that average CPP values rose in females to about age 40, and then fell slightly after age 60. Results on HNR are sparse and contradictory. In a lifespan study (age 4–93 years), Stathopoulos et al. (2011) observed that HNR rose in American English-speaking females to age 50 and fell after that. However, Fougeron et al. (2021), studying French speakers between 20–93 years, found that HNR fell to about age 50 and then remained stable. Both of these studies drew their measures from sustained productions of /a/, so the differing findings suggest that linguistic, demographic, or cultural factors may be important variables.

Measures of voice spectral tilt are rare in studies of vocal aging (cf. Schötz 2006), and the data do not lead to clear predictions. Xue and Deliyski (2001) observed that older females had relatively more energy in lower (70–1600 Hz) frequency ranges than higher ones (1600–4500 Hz) compared to younger ones. In contrast, Decoster and Debruyne (1997) found a greater balance of energy at higher frequencies in older females compared to younger ones in measures comparing H1–H2 (the amplitudes of the first and second harmonics) and E1–45 (mean energy up to 1 kHz and between 4–5 kHz). Karlsson and Hartelius (2021) recently observed complex patterns of age-related change in multiple Mel Frequency Cepstral Coefficients and their variability for sustained /a/: Relative energy increased in some narrow frequency bands, decreased in others, and showed U-shaped patterns in still others. Patterns also differed for men and women. Although considerably more work is needed to understand how spectral balance may vary with age, it does appear that aspects of spectral tilt could contribute to listener perceptions of speaker age (Schultz et al. 2021).

In sum, as with formants, some acoustic measures of laryngeal characteristics have potential explanations in terms of anatomical and physiological changes. However, linguistic factors, including prosody, should also be considered.

1.3. Research Questions and Predictions

The results so far do not provide a uniform picture of age-related changes in the use of filler particles. Based on Gerstenberg (2015), we did not expect increased use of *euh*.

That study assessed 29 LangAge speakers recorded seven years apart. The speakers of the present study were included in that corpus.

Work focusing on physical effects of aging leads to the general prediction of lower f0 in older female speakers. However, results on f0 in fillers, across languages and age, are mixed, possibly because authors have not analyzed the data based on prosodic context. Generally, we might predict that utterance-final positions have lower f0 as a function of declination (e.g., Gendrot and Schmid 2011) and utterance-final use of creaky phonation (Aare et al. 2018; Ogden 2001). Zhang (2016b) has also suggested that speakers may make glottal adjustments to conserve airflow reserves under some conditions (e.g., long utterances), such as adopting more adducted laryngeal postures. It is also possible that age interacts with prosodic patterns so that, e.g., use of utterance-final low f0 and creaky voice varies with age as a speaker-chosen stylistic pattern.

Data on the formant characteristics of fillers in French speakers are sparse. Based on previous acoustic studies of aging, we expected to find a decrease in F1 with age, but no effects of position on formants. We included F2 among our measures for completeness, but the literature did not lead to clear predictions here. We did not undertake an assessment of F3, F4 and higher formants because (a) they are sparsely represented in past work and (b) higher-frequency formants can often not be reliably detected, especially in quiet or breathy speech.

For voice quality, the documented laryngeal changes would suggest greater noise, reflecting breathier or less periodic phonation, in older speakers. Breathiness could also correspond with greater spectral tilt. Utterance-final use of creaky voice could lead to greater aperiodicity and a flatter spectrum (Childers and Wu 1991) as a function of position.

As noted above, much of the emphasis on aging in females focused on menopausal changes. Limited data for post-menopausal females suggest that aspects of the voice may continue to change, but such changes may be relatively modest (for review see Schötz 2006, chp. 4). This observation tempers our expectations for the magnitude of aging effects.

In sum, the predictions were as follows:

1. No increase in *euh* usage with age.
2. Lower f0 in utterance-final positions.
3. Lower F1 with age, but no effect of position.
4. Greater noise and steeper spectral tilt with aging; greater aperiodicity and flatter spectral tilt in utterance-final position.

2. Materials and Methods

2.1. Database

We drew on the LangAge corpus (Gerstenberg 2005–2021), which collected biographical narrative interviews with French speakers living around Orléans in the central part of France where the standard French variety dominates. The first interview series started in 2005, with various ways of approaching participants (in the street, friend of a friend, retirement homes, writing classes) so that different social milieus were included. The interviewer (third author of the paper) used open-ended questions regarding participants' memories of World War II. This protocol provided an engaging and interactive context for eliciting spontaneous speech. Participants were re-contacted in subsequent years to see if they were willing to be interviewed again. Follow-up meetings were conducted by the same interviewer revisiting the same themes.

2.2. Speakers

The current analysis includes 10 females interviewed in both 2005 and 2015/2016. Speaker ages were 63–86 years in the first session, and 74–96 years in the second. We restricted our analysis to females, given the evidence cited above that aging may have differential effects on male vs. female voices. In all cases, participants were judged to be normally-aging based on responsivity and appropriate discourse pragmatics (Gerstenberg 2011). In all cases but one (participant 037), the speakers were living independently at the

time of both the first and second recording sessions (015 had moved to a retirement home in between).

2.3. Instrumentation

The initial interviews from 2005 were recorded using a SONY Minidisk (MZ-R91) recorder and a Philips SBC ME570 condenser microphone. The later ones from 2015 used an Olympus Linear PCM recorder (LS-5) with the same Philips SBC ME570 microphone or the microphone built into the Olympus recorder. Recordings were carried out in the participants' homes (or apartment within the retirement home, in the case of speaker 037). The resulting recordings were of good acoustic quality but had the natural variations that arise in sociolinguistic fieldwork.

2.4. Utterances Analyzed

In initial assessments (Koenig et al. 2020), we used orthographic transcriptions to find frequent discourse particles, including *euh*, usually realized as [ø] or [ae], and *alors*. In the current study, we restricted the analysis to *euh*. A tier was added to the textgrids in Praat (retrieved 22 July 2020 from www.praat.org; v. 6.0.36), in which we manually marked the phonetic boundaries of *euh* with stable f0 and visible formants. Tokens were coded as being initial (ini), internal (int) or at the end (end) of an interpausal unit (IPU). We consider utterances to be demarcated by pause, so that in our usage, IPU = utterance. Criteria were as follows: Initial had a silent pause of 250 ms or longer just before *euh*; there was no pause when *euh* occurred in internal position; and a pause of minimally 250 ms followed *euh* when it occurred at the end. (When *euh* was preceded or followed by short silent intervals <250 ms, it was labeled as int). When *euh* was separated by two pauses and was an IPU on its own, the position was labelled as being between pauses (btwPaus). In some cases, *euh* was followed by a long portion of creaky voice, but no silent pause. While creaky voice may reflect a form of filled pause, we decided to label those productions of *euh* as internal, because creaky voice could not clearly be separated from the discourse particle. The 250 ms threshold was set arbitrarily, but it is similar to values used in studies of fluency (cf. de Jong and Bosker 2013), i.e., in which pause is assessed in discourse contexts. Tokens were omitted at this stage based on (a) overtalk or interfering background noise and/or (b) lack of phonation. The reason for the latter is that the analysis (see next section) requires detection of glottal cycles, i.e., f0. Examples of the four contexts are provided in Appendix A, Figure A1.

Note that we did not attempt to assess age-related changes in the duration of *euh*. In many cases, *euh* was produced as part of a continuous speech stream; when *euh* was contiguous with a vocalic element, clear acoustic boundaries were not evident. Further, in utterance-final positions *euh* productions sometimes had long breathy or creaky offsets. In those cases, offsets could be difficult to identify against background noise. Moreover, formant structure could continue in the absence of phonation. As stated above, our labeling procedure marked off the region of phonated *euh* productions that showed formant structure for F1 and F2 in the spectrogram. With these labeling criteria, it was not clear that a durational comparison across ages would be valid.

2.5. Processing

The first step was to extract the vowels in Praat using the manually delimited textgrid boundaries. Preliminary analyses indicated that very short vowels frequently did not yield measures of f0 or formants (no data, or highly unstable); thus, only tokens longer than 100 ms were extracted. These were subsequently processed using VoiceSauce [www.phonetics.ucla.edu/voicesauce; v. 1.34, last accessed 14 December 2021] with a sampling rate of 16 kHz. Initial parameter settings used Praat to estimate f0 (range 40–600 Hz, smoothing bandwidth = 10 ms) and formants (N = 4 formants, maximum formant frequency = 5500 Hz). The resulting output was reviewed to determine if a large number of productions had major discontinuities in f0 (e.g., halving or doubling) or F1. In such

cases, parameters were adjusted on a speaker- and token-specific basis and VoiceSauce was re-run. These final VoiceSauce files were reviewed with reference to the waveforms and spectrograms, and occasional spurious values of f0 or F1 were trimmed out in MATLAB (2018). Final exclusions were made in cases where (a) VoiceSauce failed to complete the analysis, (b) the resulting values of f0 or F1 were greatly at odds with the spectrogram or the percept, and/or (c) a region of stable f0 for most of the token could not be determined based on the harmonic structure in a narrowband spectrogram. The complete dataset with 10 speakers, 2 timeframes, and 4 prosodic positions consisted of 1452 samples.

2.6. Measures

Of the numerous output measures provided by VoiceSauce, Table 1 depicts the ones chosen for statistical analysis, focusing specifically on voice quality parameters, because they have been the least investigated.

Table 1. Overview of the investigated speech characteristics and associated acoustic parameters. Measures of voice quality are grouped thematically.

System Level	Acoustic Parameters
Supralaryngeal resonances	•First formant (F1) •Second formant (F2)
Phonation	•Fundamental frequency (f0)
Voice quality	Low-frequency spectral tilt (up to about 1000 Hz): •H1H2 (amplitudes of first and second harmonics) •H2H4 (amplitudes of second and fourth harmonics) •H1A1 (amplitudes of first harmonic and F1)
	Higher-frequency spectral tilt (up to about 5000 Hz): •H1A2 (amplitudes of first harmonic and F2) •H1A3 (amplitudes of first harmonic and F3) •H42K (amplitudes of fourth harmonic and that at ca. 2 kHz) •H2K5K (amplitudes of second harmonic and that at ca. 5 kHz)
	Noise, aperiodicity: •HNR35 •CPP

VoiceSauce provides four measures of harmonics to noise ratio, with overlapping frequency ranges (0–500, 0–1500, 0–2500, and 0–3500 Hz). Before settling on the final dependent measures, we obtained a correlation matrix for all variables. This analysis indicated, not surprisingly, that the four HNR measures were highly correlated, with r-values 0.85–0.99. The lowest of these r-values was between the most distant ranges, i.e., HNR05 and HNR35. For final analysis, we chose HNR35 alone. Other correlations greater than r = 0.70 were among the input measures A1, A2, A3, H1, H2, H2k, i.e., among the amplitudes of the first three formants (A1–A3) and harmonic amplitudes H1, H2, and that closest to 2000 Hz. None of the spectral tilt measures themselves showed such high intercorrelations.

All spectral tilt measures were corrected to account for differences in vowel quality (Hanson 1997; Iseli et al. 2007). All VoiceSauce outputs measure over time (here, every 10 ms). For this analysis, we did not explore variation over time but entered average values for each measure and token into the statistical analyses.

2.7. Statistics and Exclusions

All statistical analyses were carried out using R (2021), version 4.1.0. Several linear mixed-effects models were run using the packages *lme4* (Bates et al. 2015) and *lmerTest* (Kuznetsova et al. 2017). The *emmeans* package (Lenth 2021) was used for posthoc analyses, with a confidence level of 0.95 (alpha level = 0.05), an adjustment of p-values using the Tukey method, and the Kenward Roger method to calculate degrees of freedom.

All graphical explorations were carried out with *ggplot2* (Wickham 2016) and *ggpubr* (Kassambara 2020).

Separate linear mixed-effects models were run for all measures described in Table 1, which were treated as dependent variables. We took TIME (2005 vs. 2015, reference level 2005) and POSITION (ini vs. int vs. end, reference level ini) as independent factors and allowed for their interaction. Since *euh* occurring between pauses was rare or nonexistent for some speakers, it was excluded from the analysis. Speaker specific slopes and intercepts for TIME by speaker and POSITION by speaker served as random effects. When post-hoc analyses were called for, we did not assess comparisons involving different POSITIONS across TIME (e.g., comparing utterance-initial position in 2005 with utterance-final position in 2015) because we did not consider them as meaningful.

The composition of the final dataset, after the exclusion of btwPaus tokens, is provided in Table 2, showing the quantity of data for each speaker and time. A more detailed breakdown across speakers is provided in Appendix A, Table A1.

Table 2. Breakdown of productions, as a percentage of 1383 tokens, for each speaker and time. Positional contexts for *euh* were initial, internal, and end.

Speakers	11	15	16	19	26	29	37	46	49	50	Total
2005	3.5	4.4	7.6	4.5	6.1	5.5	5.4	5.2	6.9	4.5	53.6
2015	1.4	2.6	4.6	1.9	5.5	6.8	5.2	6.7	7.4	4.3	46.4
Total	4.8	7.0	12.2	6.4	11.6	12.3	10.6	11.9	14.4	8.8	

3. Results

We begin with reporting the results for *euh* frequency, and then proceed to the acoustics (formants, f0, and voice quality). A thematic summary of the acoustic results is provided in Table 3 below. The violin plots for f0, formants, and voice quality (Sections 3.2–3.4) show group data; speaker-specific data are displayed in the Appendix A (Figures A1–A8).

Table 3. Summary of significant effects of TIME (2005 vs. 2015) and POSITION (initial, internal, end), along with interactions.

Parameter	TIME	POSITION	Interaction
F2		ini > int, end	ini > int, end in 2015
f0		ini > end	ini, int > end in 2005
H1A1	2005 > 2015		ini, int < end in 2015
H42K		ini > int	ini > end, ini $\geq$ int (p = 0.06) in 2015
H2KH5K		ini < int, end	ini < int, end in 2015
HNR35	2005 < 2015		
CPP		ini > end	ini, int > end in 2005

3.1. Frequency of euh

We did not find evidence for increased usage of *euh* in the older participants. Instead, the tendency in our corpus was a decrease of normalized *euh* frequency: Median *euh* usage was 30 occurrences per 1000 tokens in 2005 vs. 25/1000 in 2015. There was a higher standard deviation in 2015 (10 *euh* per 1000 words in 2005 vs. 17 in 2015). Since the frequency of *euh* per 1000 tokens was normally distributed, we performed a paired t-test and additionally, due to the small sample size, a Wilcoxon rank sum exact test. With t = 0.24, df = 9, p = 0.81 (*t*-test) and W = 60 and p = 0.48 (Wilcoxon), both tests revealed no significant difference in the frequency of *euh* between the two time points.

3.2. Formants

Results for the two formants revealed only significant effects for F2 (Figure 1A; see also Appendix A Figure A2), with no effects for F1. Speakers produced significantly lower F2 values, i.e., a more centralized or backed *euh*, for the internal position (intercept = 1739 Hz, estimate = −63.76 Hz, SE = 27.63, t = −2.31, *p* = 0.0442) and at the end of an IPU (end estimate: −62.15 Hz, SE = 23.67, t = −2.63, *p* = 0.0235) in comparison to the initial position. The effect of TIME was not significant. A posthoc analysis revealed an interaction, whereby the positional effect on F2 was found for the later recording sessions, but not the 2005 series (2015 ini vs. 2015 int, estimate = 103 Hz, SE = 28.5, t = 3.63, *p* = 0.0300; 2015 ini vs. 2015 end, estimate = 85 Hz, SE = 24.0, t = 3.5, *p* = 0.0325).
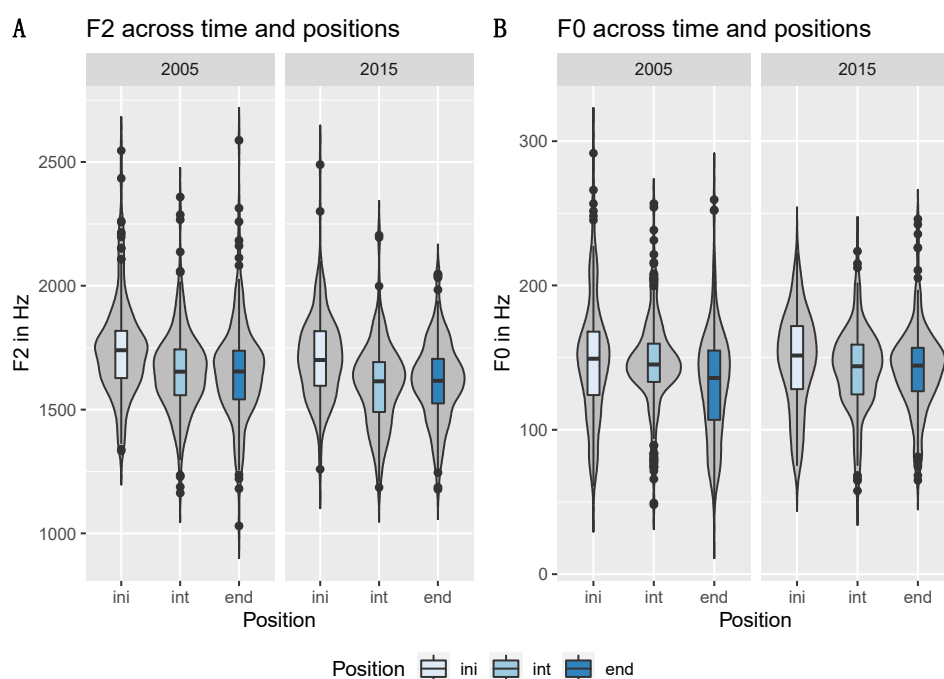
Figure 1. Violin plots showing rotated kernel density and boxplots with median values (black horizontal line) for (**A**) F2 in Hz and (**B**) f0 in Hz. Data are split by TIME of recording (2005 and 2015) and POSITION (initial, internal, end) of the filler particle.

3.3. Fundamental Frequency

Fundamental frequency did not show a main effect of TIME, but an effect of POSITION was found (Figure 1B and Appendix A Figure A3). F0 was lower when *euh* occurred at the end of an interpausal unit in comparison to the initial position (intercept = 145 Hz, estimate = −11.31 Hz, SE = 2.98, t = −3.78, *p* < 0.001). The posthoc analysis revealed an interaction between the two factors. At the earlier recordings, *euh* at the end of an IPU had a lower F0 than in internal (2005 int vs. 2005 end, estimate = 11.39 Hz, SE = 2.73, t = 4.18, *p* = 0.0039) and initial position (2005 ini vs. 2005 end, estimate: 11.31 Hz, SE = 3.06, t = 3.67, *p* = 0.0105). In this case, 10 years later no difference among prosodic positions was found.

3.4. Voice Quality Parameters

The low frequency spectral tilt parameters H1H2 and H2H4 were not affected by TIME or POSITION, nor their interaction. For H1A1 (Figure 2A and Appendix A Figure A4) we found a main effect of TIME, with lower average values in 2015 than 2005 (intercept: 1.23, estimate: −2.5, SE = 1, t = −2.54, *p* = 0.0239). The posthoc analysis indicated that in the later recording, speakers had lower values in initial and internal position than at the end (2015 ini vs. 2015 end, estimate = −2.2, SE = 0.61, t = −3.6, *p* = 0.0198; 2015 int vs. 2015 end, estimate = −2.0, SE = 0.6, t = −3.2, *p* = 0.0439).
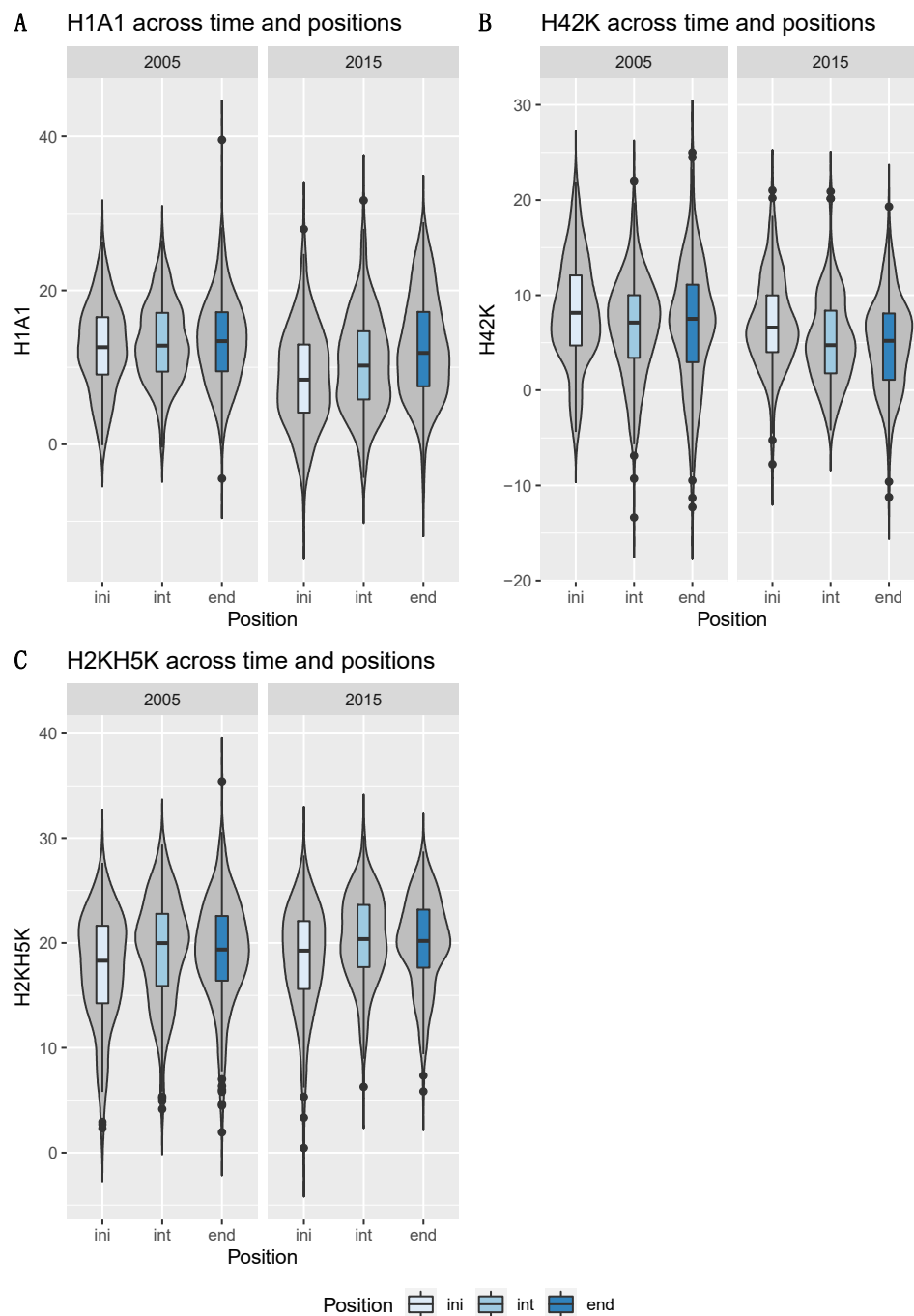
Figure 2. Violin plots showing rotated kernel density and boxplots with median values (black horizontal line) for spectral tilt parameters ((**A**): H1A1, (**B**): H42K, (**C**): H2KH5K). Data are split by TIME of recording (2005 and 2015) and POSITION (initial, internal, end) of the filler particle.

The higher-frequency tilt measures, H1A2 and H1A3, did not differ according to TIME or POSITION, or their interaction. The parameter H42K (Figure 2B and Appendix A Figure A5) showed an effect for POSITION, but not TIME: Values were lower for the internal position in comparison to initial (intercept: 7.96, estimate: -1.13, SE = 0.53, t = -2.14, p = 0.0447). In the posthoc analysis, an effect emerged only for 2015 between the initial and final position (2015 ini vs. 2015 end, estimate: 1.6, SE = 0.50, t = 3.17, p = 0.0433). The 2015 comparison of initial versus internal almost reached significance (p = 0.06).

The H2KH5K parameter (Figure 2C and Appendix A Figure A6) demonstrated no effects for TIME but a main effect for POSITION, with lower values in initial position than

in the other two positions (intercept = 9.56; int estimate = 1.17, SE = 0.44, t = 2.68, *p* = 0.01147; end estimate = 1.21, SE = 0.44, t = 2.75, *p* = 0.00702). The posthoc analyses revealed again that these differences were only found at the later recording (2015 ini vs. 2015 int, estimate −1.81, SE = 0.48, t = −3.74, *p* = 0.0106; 2015 ini vs. 2015 end, estimate = −1.47, SE = 0.46, t = −3.21, *p* = 0.0386).

HNR35 (Figure 3A and Appendix A Figure A7) was the only parameter with an effect of TIME only, with no interaction with POSITION (intercept = 37.85, estimate = 5.95, SE = 2.29, t = 2.6, *p* = 0.0242). HNR values were higher in the later recording sessions.
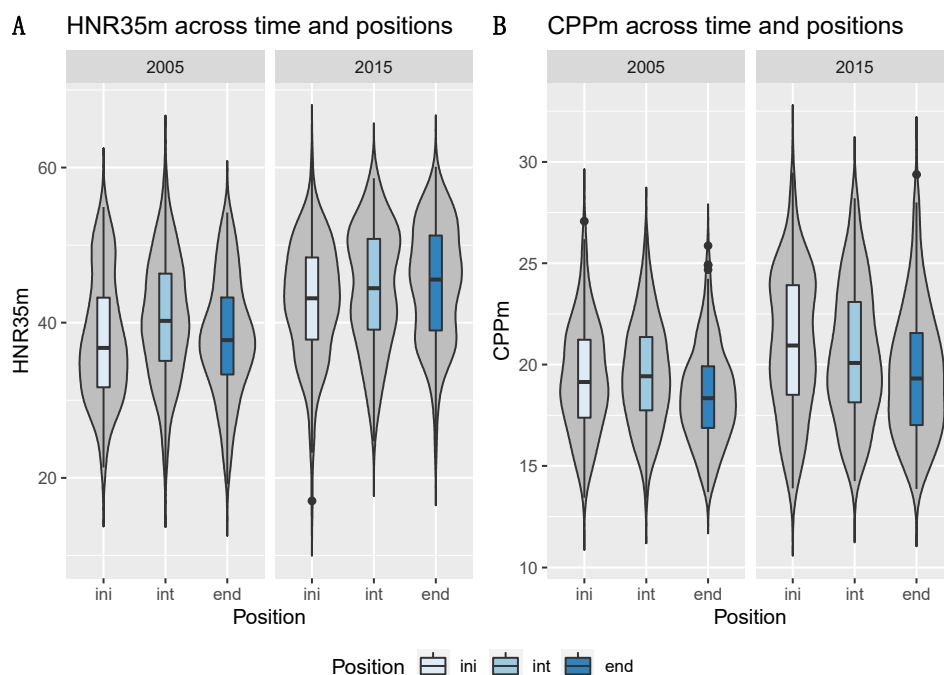
Figure 3. Violin plots showing rotated kernel density and boxplots with median values (black horizontal line) for noise measures ((**A**): HNR35, (**B**): CPP). Data are split by TIME of recording (2005 and 2015) and POSITION (initial, internal, end) of the filler particle.

Finally, findings for Cepstral Peak Prominence (CPPm) (Figure 3B, Appendix A Figure A8) showed a main effect of POSITION, with higher values in the initial position compared to final (intercept = 19.66, estimate = −1.15, SE = 0.35, t = −3.32, *p* = 0.00406). In the posthoc analysis, positional effects held only for the 2005 recording (2005 ini vs. 2005 end, estimate = 1.15, SE = 0.35, t = 3.28, *p* = 0.0436; 2005 int vs. 2005 end, estimate = 0.97, SE = 0.23, t = 4.23, *p* = 0.0028).

Table 3 summarizes the acoustical results. The most robust main effects based on the *p*-values were found for positional effects on f0 (*p* < 0.001) and CPP (*p* = 0.00406).

4. Discussion

The results of this work can be summarized as follows: (a) *euh* frequency did not increase with age, and (b) effects of utterance position were more common than effects of aging on acoustic parameters. Aging effects were found only for measures of voice quality.

4.1. Frequency of euh

As expected, based on prior work using the LangAge corpus (Gerstenberg 2015), we did not detect an age-related increase in *euh* usage. This is comparable to other work recording either no change (Keszler and Bóna 2019; Searl et al. 2002) or even a decrease (Gall 2019; Gósy and Silber-Varod 2021; Maxim et al. 1994; Taschenberger et al. 2019) in the use of some FPs with age. Given that FPs have multiple discourse functions, and behave differently within speakers (Bortfeld et al. 2001; Horton et al. 2010), we suggest that it is

inappropriate to make blanket predictions about how discourse particles, as a group, may change as a function of age. Based on our results for older speakers, and the participants of Searl et al. (2002), who were 70 years and older, the usage patterns of some filler particles could be relatively stable at later ages.

4.2. Effects of Age and Position

4.2.1. Age

Among all measures, the only two with main effects of TIME reflected voice quality: Low–mid frequency spectral tilt (H1A1) and noise in frequencies up to 3500 Hz (HNR35). For H1A1, the interpretation is complicated by interactions with POSITION (see next section). This left only HNR35 as the one clear acoustic age marker in this dataset. The direction of the difference was unexpected, however: The ratio of harmonic energy relative to noise increased with age, i.e., there was relatively less noise. This is inconsistent with reports of bowing, glottal gaps, and other forms of laryngeal insufficiency at advanced ages, giving rise to breathier voice qualities. One possible explanation is that edema in older females could increase vocal fold contact (cf. Biever and Bless 1989; Honjo and Isshiki 1980; Pontes et al. 2005; cf. also Higgins and Saxman 1991; Kahane 1987b) leading to reduced noise (Zhang 2016a).

It could also be that older speakers use more compressed glottal setting to compensate for a reduced respiratory drive (cf. Zhang 2016b). Although positional effects did not reach significance, qualitatively HNR35 was lowest in utterance-initial position, when lung volumes, on average, should be higher. Finally, speakers could adopt different voice qualities as a stylistic feature as they age. A traditional finding in the literature is that anatomical differences should, on average, lead to breathier voice qualities in females than males (e.g., Hanson 1997); it is also clear, however, that sociocultural influences can override such tendencies (e.g., Wagner and Braun 2003; Wolk et al. 2012; Yuasa 2010).

The lack of significant age effects for f0 and formants was surprising. As described in the Introduction, changes in these parameters have been observed in several studies of aging. However, the analyses of Schötz (2006) for speakers 20–90 years of age showed that f0 and F1 in females decrease from young adulthood to middle age, with more stability later on (cf. also Fougeron et al. 2021; Stathopoulos et al. 2011). That is, lifespan studies of f0 in females may arrive at different results depending on which decades are studied. Finally, as discussed earlier (Section 1.1.3), results obtained for FPs may also differ from those obtained for other kinds of speech materials.

4.2.2. Utterance Position: Main Effects and Interactions

The effect of POSITION on f0 was as expected, with lower values in the final compared to the initial position. This decrease could reflect reduced respiratory supply, declination, and/or greater use of creaky voice utterance-finally. On the other hand, formant frequency changes as a function of position were not expected. The greater centralization (lower F2) of *euh* in later positions could reflect a form of supralaryngeal declination (e.g., Tabain 2003; Vayra and Fowler 1992). In general, centralization may be more typical in connected speech than in other speech tasks such as reading.

For the voice quality measures, Cepstral Peak Prominence (CPP) revealed a positional effect that, in post-hoc tests, emerged only for the 2005 recording sessions. This could indicate (cf. Schultz et al. 2021) that utterance-final creak was more extensive or consistent in earlier recording sessions. This would accord with the patterns found for f0.

Three spectral tilt measures had either main effects of POSITION (H42K, H2KH5K) or had an interaction with POSITION (H1A1). In all cases, significant effects emerged in the 2015 recording sessions only. In a post-hoc assessment, we explored the input parameters to these various tilt measures to understand these effects. In the case of H1A1, the effect mainly arose from changes in the amplitude of F1, i.e., H1 amplitude did not change much. For H42K and H2KH5K, inspection of the inputs showed in both cases that lower amplitude at 2 kHz changed the ratio. Thus, in 2015, frequencies in the range of F1 and at

2 kHz showed reduced amplitudes in later utterance positions (i.e., int and end compared to init). Interestingly, summing over POSITION, the H1A1 age difference mainly reflected changes in H1 rather than A1.

The finding of more voice quality changes as a function of POSITION in 2015 could indicate more laryngeal adjustments, possibly as a mechanism of airflow conservation, at the later ages. More generally, these interactions between POSITION and TIME suggest that prosodic functions may mediate the effects of physical aging (cf. Cole 2015 on the observation that prosody interacts with multiple other aspects of speech and language). Lifespan studies of speech characteristics accordingly should recognize possible interplays among physical, linguistic, and pragmatic effects during aging (cf. Gerstenberg 2020).

4.3. Limitations and Future Work

Although our study contributes to the relatively small body of longitudinal studies of aging, our sample size was not large. Further work is needed to determine to what extent these results can be generalized, both within the demographic studied here, and to different groups of older speakers across cultures.

The LangAge project, similar to all studies employing volunteer participants, probably tended to draw the interest of people who were at least as healthy as average for their cohort (cf. Eichhorn et al. 2018; Michel et al. 1987). Some data suggest that effects of 'aging' may be considerably more pronounced in persons in poorer physical condition (Ramig and Ringel 1983). Further, our results on French may not hold for speakers of other languages. It is not clear to what extent the nature of vocal aging may vary with social and cultural factors (cf. Michel et al. 1987; Ringel and Chodzko-Zajko 1987).

The extensive variation across prosodic positions and speakers may have masked some possible differences. For example, although CPP did not show a significant age effect, qualitatively (see Figure 3B) values were higher in the later recording sessions, in parallel with HNR035 (Figure 3A), but the CPP data were quite variable, especially in 2015. Indeed, most of our statistical effects were weak.

Finally, although we included speaker-specific slopes in all analyses, we have not extensively explored speaker differences, which could reveal subtypes of aging. Considerable inter-individual variation has been observed in anatomical and physiological aging (Hammond et al. 2000; Hirano et al. 1983; Kahane 1988; Pressman and Kelemen 1955; Turk and Hogg 1993). The same presumably holds for other aspects of aging.

5. Conclusions

Our results for the frequency of the filler *euh* in spoken French are not consistent with the notion that hesitation phenomena are generally increased in older speakers. They do point to the need to assess FPs in terms of prosodic and pragmatic contexts.

Formants, f0, and aperiodicity have been widely reported in the literature on vocal aging, with less attention given to spectral measures of voice quality. Our results suggest that such measures may provide a rich source of information on vocal changes with aging. Given the wide range of voice qualities in female speakers reported by Hanson (1997), and the increasing cross-speaker variability that arises with aging (Ringel and Chodzko-Zajko 1987), speaker specific analyses may be needed to draw clear conclusions from such data.

For understandable reasons, past acoustic studies of vocal aging tended to use restricted speech materials (sustained vowels and/or reading). Results from such controlled and somewhat atypical speech tasks may not carry over well to spontaneous speech. The complex patterns observed in our data, with numerous interactions between TIME and POSITION, suggest that a full lifespan perspective on speech characteristics should consider interactions among physical, sociolinguistic, pragmatic, and cultural factors. We acknowledge the statistical challenges of such work. Case studies and/or large corpus analyses as well as combining the factors in principal component analyses or the like could contribute to a better understanding of the multidimensional nature of how human aging is reflected in the speech signal.

Appendix A

Figure A1. *Cont.*

Figure A1. Examples of *euh* in the four labeled contexts. From top to bottom: initial, internal, end, and between pause.

Figure A2. Speaker-specific data for F2.

Figure A3. Speaker-specific data for f0.

Figure A4. Speaker-specific data for H1A1.

Figure A5. Speaker-specific data for H42K.

Figure A6. Speaker-specific data for H2KH5K.

Figure A7. Speaker-specific data for HNR35Km.

Figure A8. Speaker-specific data for CPPm.

Table A1. The data on total *euh* quantity per speaker (columns 3–5) were drawn from the LangAge corpus. Columns 6–11 show the data analyzed here, including the breakdown across positional contexts. Since th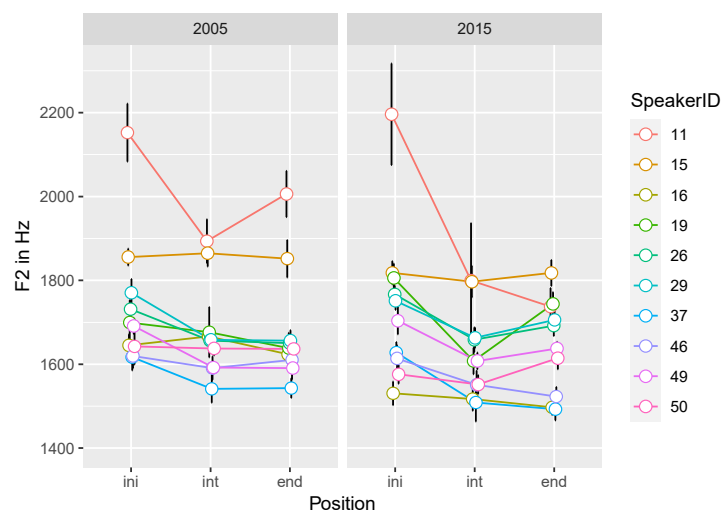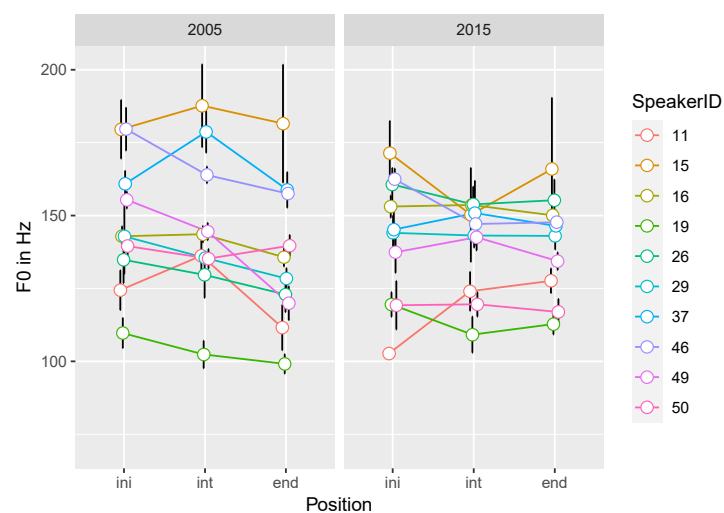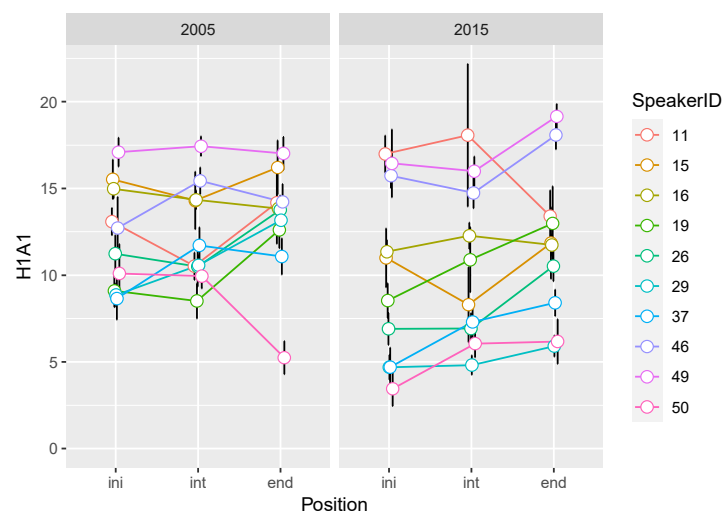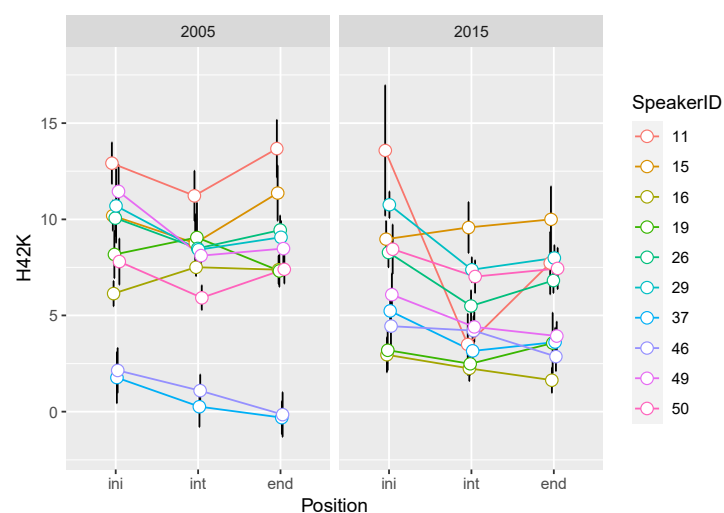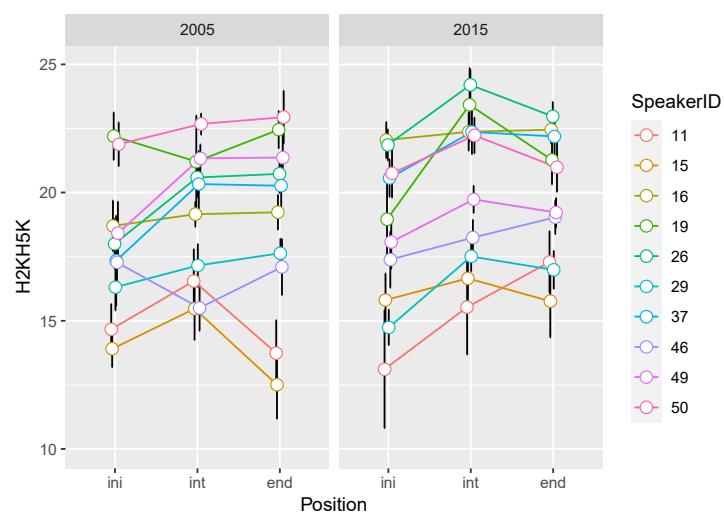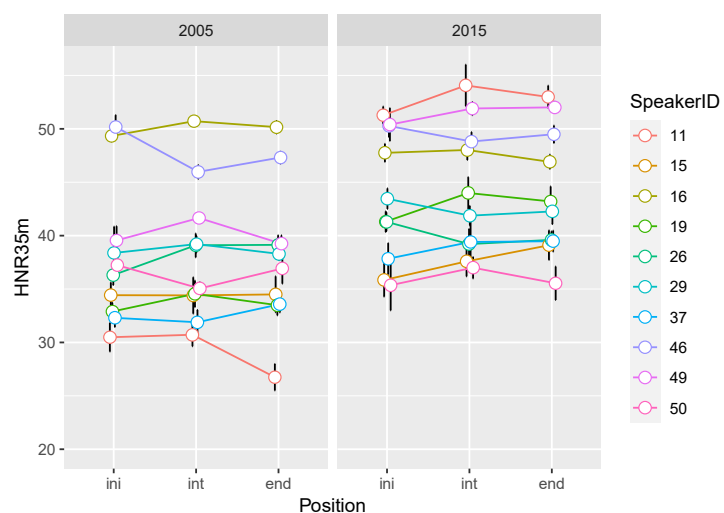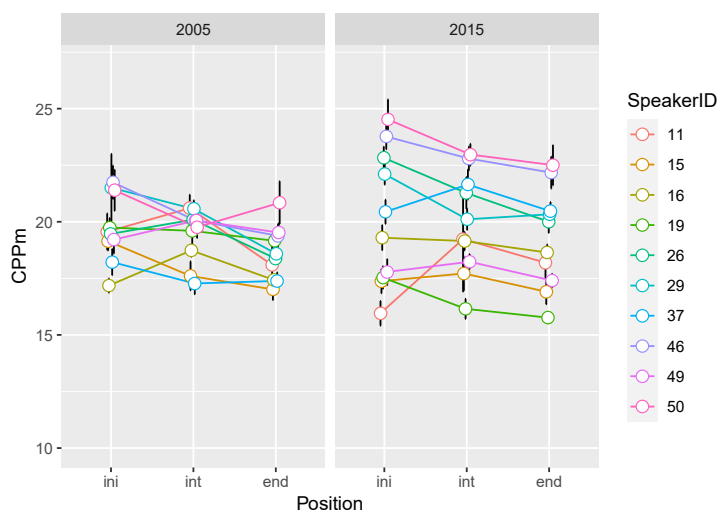e between-Pause context was small and not balanced across speakers, the number of analyzed *euh* includes the other three contexts (ini, init, end) only. The data show extensive speaker differences in *euh* usage which can be explored in future work.

Sp	Age	No. Tokens	No. All *euh*	No. *euh*/1000 Tokens	No. Analyzed *euh*	Pct (%) Analyzed *euh*	No. btwPaus	No. ini	No. int	No. End
11	86	8990	100	11.1	48	48	1	12	19	17
15	78	5923	106	17.9	61	58	6	35	15	11
16	71	6068	252	41.5	105	42	4	25	57	23
19	76	4545	99	21.8	62	63	3	16	12	34
26	78	7369	183	24.8	85	46	3	31	20	34
29	67	5059	148	29.3	76	51		7	31	38
37	79	5438	170	31.3	74	44	2	17	22	35
46	74	5931	221	37.3	72	33	4	11	37	24
49	72	9865	304	30.8	96	32		12	58	26
50	63	6018	247	41	62	25		12	43	7
11	96	9817	87	8.9	19	22		4	4	11
15	88	5725	61	10.7	36	59	6	16	11	9
16	81	3914	246	62.9	64	26	4	12	27	25
19	86	7134	88	12.3	26	30	1	15	5	6
26	88	6429	113	17.6	76	67	6	29	21	26
29	77	5571	175	31.4	94	54		33	29	32
37	89	6307	174	27.6	72	41	14	24	12	36
46	84	4259	195	45.8	92	47	9	25	34	33
49	82	6857	156	22.8	103	66	5	10	40	53
50	74	4051	115	28.4	60	52	1	10	35	15

References

Aare, Kätlin, Pärtel Lippus, Marcin Włodarczak, and Mattias Heldner. 2018. Creak in the Respiratory Cycle. Paper presented at the Interspeech, Hyderabad, India, September 2–6; pp. 1408–12. [CrossRef]

Albuquerque, Luciana, Catarina Oliveira, António Teixeira, Pedro Sa-Couto, and Daniela Figueiredo. 2020. A comprehensive analysis of age and gender effects in European Portuguese oral vowels. *Journal of Voice*. in press. [CrossRef]

Bates, Douglas, Martin Mächler, Benjamin M. Bolker, and Steven Walker. 2015. Fitting linear mixed-effects models using lme4. *Journal of Statistical Software* 67: 1–48. [CrossRef]

Belz, Malte. 2020. Acoustic Vowel Quality of Filler Particles in German. Paper presented at the Laughter and Other Non-Verbal Vocalisations Workshop, Bielefeld, Germany, October 5; pp. 7–10. [CrossRef]

Belz, Malte. 2021. *Die Phonetik von äh und ähm. Akustische Variation von Füllpartikeln im Deutschen*. Stuttgart: Metzler.

Benjamin, Barbaranne J. 1981. Frequency variability in the aged voice. *Journal of Gerontology* 36: 722–26. [CrossRef]

Biever, Dawn M., and Diane M. Bless. 1989. Vibratory characteristics of the vocal folds in young adult and geriatric women. *Journal of Voice* 3: 120–31. [CrossRef]

Bolly, Catherine T., George Christodoulides, and Anne Catherine Simon. 2016. Disfluences et vieillissement langagier. De la base de données Valibel aux corpus outillés en français parlé. *Corpus* 15: 1–13. [CrossRef]

Bortfeld, Heather, Silvia D. Leon, Jonathan E. Bloom, Michael F. Schober, and Susan E. Brennan. 2001. Disfluency rates in conversation: Effects of age, relationship, topic, role, and gender. *Language and Speech* 44: 123–47. [CrossRef]

Brown, W. S., Richard J. Morris, and John F. Michel. 1989. Vocal jitter in young adult and aged female voices. *Journal of Voice* 3: 113–19. [CrossRef]

Candea, Maria, Ioana Vasilescu, and Martine Adda-Decker. 2005. Inter- and Intra-Language Acoustic Analysis of Autonomous Fillers. Paper presented at the Disfluency in Spontaneous Speech Workshop, Aix-en-Provence, France, September 10–12; pp. 47–51. Available online: https://www.isca-speech.org/archive_open/diss_05/dis5_047.html (accessed on 15 December 2021).

Candea, Maria. 2000. Contribution à l'Etude des Pauses Silencieuses et des Phénomènes dits 'd'Hésitation' en Français Oral Spontané. Ph.D. dissertation, Université Paris, Paris, France.

Childers, Donald G., and Ke Wu. 1991. Gender recognition from speech. Part II: Fine analysis. *The Journal of the Acoustical Society of America* 90: 1841–56. [CrossRef]

Cole, Jennifer. 2015. Prosody in context: A review. *Language, Cognition and Neuroscience* 30: 1–31. [CrossRef]

de Jong, Nivja H., and Hans Rutger Bosker. 2013. Choosing a Threshold for Silent Pauses to Measure Second Language Fluency. Paper presented at the Disfluency in Spontaneous Speech Workshop, Stockholm, Sweden, August 21–23; pp. 17–20. Available online: http://hdl.handle.net/11858/00-001M-0000-0015-0FB8-8 (accessed on 15 December 2021).

Decoster, Wivine, and Frans Debruyne. 1997. Changes in spectral measures and voice-onset time with age: A cross-sectional and a longitudinal study. *Folia Phoniatrica et Logopaedica* 49: 269–80. [CrossRef] [PubMed]

Dehquan, Ali, Ronald C. Scherer, Gholamali Dashti, Alireza Ansari-Moghaddam, and Sepideh Fanaie. 2012. The effects of aging on acoustic parameters of voice. *Folia Phoniatrica et Logopaedica* 64: 265–70. [CrossRef] [PubMed]

Duez, Danielle. 2001. Acoustico-phonetic Characteristics of Filled Pauses in Spontaneous French Speech: Preliminary Results. Paper presented at the Disfluency in Spontaneous Speech Workshop, Edinburgh, Scotland, UK, August 29–31; pp. 41–44. Available online: https://www.isca-speech.org/archive_open/diss_01/dis1_041.html (accessed on 15 December 2021).

Eichhorn, Julie T., Raymond D. Kent, Diane Austin, and Houri K. Vorperian. 2018. Effects of aging on vocal fundamental frequency and vowel formants in men and women. *Journal of Voice* 32: 644.e1–644.e9. [CrossRef]

Endres, Werner, W. Bambach, and G. Flösser. 1971. Voice spectrograms as a function of age, voice disguise, and voice imitation. *The Journal of the Acoustical Society of America* 49: 1842–48. [CrossRef]

Estenne, Marc, Jean-Claude Yernault, and André de Troyer. 1985. Rig cage and diaphragm-abdomen compliance in humans: Effects of age and posture. *Journal of Applied Psychology* 59: 1842–48. [CrossRef]

Ferrand, Carole T. 2002. Harmonics-to-Noise Ratio: An index of vocal aging. *Journal of Voice* 16: 480–87. [CrossRef]

Ferreira, Fernanda, and Karl G. D. Bailey. 2004. Disfluencies and human language comprehension. *Trends in Cognitive Sciences* 8: 231–37. [CrossRef] [PubMed]

Fougeron, Cécile, Fanny Guitard-Ivent, and Véronique Delvaux. 2021. Multi-dimensional variation in adult speech as a function of age. *Languages* 6: 176. [CrossRef]

Frank, N. Robert, Jere Mead, and Benjamin G. Ferris Jr. 1957. The mechanical behavior of the lungs in healthy elderly persons. *The Journal of Clinical Investigation* 36: 1680–87. [CrossRef]

Gall, Juliane. 2019. *Disfluencies im Laufe des höheren Lebensalters. Eine individuenzentrierte Längsschnittstudie*. Göttingen: Verlag für Gesprächsforschung.

Gendrot, Cédric, and Carolin Schmid. 2011. F0 Declination in French: Broadcast News versus Spontaneous Speech. Paper presented at the Nijmegen Workshop in Production and Comprehension of Conversational Speech, Nijmegen, The Netherlands, December 12–13; pp. 15–17. Available online: https://hal.archives-ouvertes.fr/halshs-00691494/ (accessed on 15 December 2021).

Gerstenberg, Annette. 2005–2021. *LangAge Collection of Biographical Interviews. Potsdam: Department of Romance Studies*. Available online: www.langage-corpora.org (accessed on 15 December 2021).

Gerstenberg, Annette. 2011. *Generation und Sprachprofile im höheren Lebensalter: Untersuchungen zum Französischen auf der Basis eines Korpus biographischer Interviews*. Frankfurt am Main: Vittorio Klostermann.

Gerstenberg, Annette. 2015. A sociolinguistic perspective on vocabulary richness in a seven-year comparison of older adults. In *Language Development: The Lifespan Perspective*. Edited by Annette Gerstenberg and Anja Voeste. Amsterdam: Benjamins, pp. 109–27. [CrossRef]

Gerstenberg, Annette. 2020. Pragmatic development in the (middle and) later stages of life. In *Developmental and Clinical Pragmatics (HOPS 13)*. Edited by Klaus P. Schneider and Elly Ifantidou. Berlin and Boston: De Gruyter Mouton, pp. 209–34. [CrossRef]

Gerstenberg, Annette, Susanne Fuchs, Julie Kairet, Johannes Schroeder, and Claudia Frankenberg. 2018. A Cross-linguistic, Longitudinal Case Study of Pauses and Interpausal Units in Spontaneous Speech Corpora of Older Speakers of German and French. Paper presented at the 9th International Conference on Speech Prosody, Poznań, Poland, June 13–16; pp. 211–15.

Goozée, Justine V., Bruce E. Murdoch, Deborah G. Theodoros, and Elizabeth C. Thompson. 1998. The effects of age and gender on laryngeal aerodynamics. *International Journal of Language and Communication Disorders* 33: 221–38. [CrossRef]

Gósy, Mária, and Vered Silber-Varod. 2021. Attached Filled Pauses: Occurrences and Durations. Paper presented at the Disfluency in Spontaneous Speech Workshop, Paris, France, August 25–26; pp. 71–76.

Gósy, Mária, Judit Bóna, András Beke, and Viktória Horváth. 2014. Phonetic Characteristics of Filled Pauses: The Effects of Speakers' Age. Paper presented at the 10th International Seminar on Speech Production, Köln, Germany, May 5–8; pp. 150–53. Available online: http://www.issp2014.uni-koeln.de/wp-content/uploads/2014/Proceedings_ISSP_revised.pdf (accessed on 22 November 2021).

Grosman, Iulia, Anne Catherine Simon, and Liesbeth Degand. 2018. Variation de la durée des pauses silencieuses: Impact de la syntaxe, du style de parole et des disfluences. *Langages* 111: 13–40. [CrossRef]

Hammond, Thomas H., Steven D. Gray, John Butler, Ruixia Zhou, and Elizabeth Hammond. 1998. Age-and gender-related elastin distribution changes in human vocal folds. *Otolaryngology—Head and Neck Surgery* 4: 314–22. [CrossRef]

Hammond, Thomas H., Steven D. Gray, and John E. Butler. 2000. Age- and gender-related collagen distribution in human vocal folds. *Annals of Otology, Rhinology & Laryngology* 109: 913–20. [CrossRef]

Hanson, Helen M. 1997. Glottal characteristics of female speakers: Acoustic correlates. *The Journal of the Acoustical Society of America* 101: 466–81. [CrossRef]

Harrington, Jonathan, Sallyanne Palethorpe, and Catherine I. Watson. 2000. Does the Queen speak the Queen's English? *Nature* 408: 927–28. [CrossRef]

Higgins, Maureen B., and John H. Saxman. 1991. A comparison of selected phonatory behaviors of healthy aged and young adults. *Journal of Speech and Hearing Research* 34: 1000–10. [CrossRef] [PubMed]

Hirano, Minoru, Shigejiro Kurita, and Teruyuki Nakashima. 1983. Growth, development, and aging of human vocal folds. In *Vocal Fold Physiology: Contemporary Research and Clinical Issues*. Edited by Diane M. Bless and James H. Abbs. San Diego: College-Hill Press, pp. 22–43.

Hoit, Jeannette D., and Thomas J. Hixon. 1987. Age and speech breathing. *Journal of Speech and Hearing Research* 30: 351–66. [CrossRef] [PubMed]

Honjo, Iwao, and Nobuhiko Isshiki. 1980. Laryngoscopic and voice characteristics of aged persons. *Archives of Otolaryngology* 106: 149–50. [CrossRef] [PubMed]

Horton, William S., Daniel H. Spieler, and Elizabeth Shriberg. 2010. A corpus analysis of patterns of age-related change in conversational speech. *Psychology and Aging* 25: 708–13. [CrossRef]

Iseli, Markus, Yen-Liang Shue, and Abeer Alwan. 2007. Age, sex, and vowel dependencies of acoustic measures related to the voice source. *The Journal of the Acoustical Society of America* 121: 2283–95. [CrossRef] [PubMed]

Israel, Harry. 1973. Age factor and the pattern of change in craniofacial structures. *American Journal of Physical Anthropology* 39: 111–28. [CrossRef] [PubMed]

Kahane, Joel C. 1987a. Connective tissue changes in the larynx and their effects on voice. *Journal of Voice* 1: 27–30. [CrossRef]

Kahane, Joel C. 1987b. Developmental changes in the articular cartilage of the human cricoarytenoid joint. In *Laryngeal Function in Phonation and Respiration*. Edited by Thomas Baer, Clarence Sasaki and Katherine S. Harris. Boston: Little, Brown and Company, pp. 14–28.

Kahane, Joel C. 1988. Age related changes in the human cricoarytenoid joint. In *Vocal Physiology: Voice Production, Mechanisms and Functions*. Edited by Osamu Fujimura. New York: Raven Press, pp. 145–57.

Karlsson, Fredrik, and Lena Hartelius. 2021. On the primary influences of age on articulation and phonation in maximum performance tasks. *Languages* 6: 174. [CrossRef]

Karpiński, Maciej. 2013. Acoustic features of filled pauses in Polish task-oriented dialogues. *Archives of Acoustics* 38: 63–73. [CrossRef]

Kassambara, Alboukadel. 2020. ggpubr: 'ggplot2' Based Publication Ready Plots, Version 0.4.0, 2020-06-27. R-Project: Package ggplot2. Available online: https://cran.r-project.org/web/packages/ggpubr/index.html (accessed on 15 December 2021).

Keszler, Borbála, and Judit Bóna. 2019. Pausing and Disfluencies in Elderly Speech: Longitudinal Case Studies. Paper presented at the 9th Workshop on Disfluency in Spontaneous Speech, Budapest, Hungary, September 12–13; pp. 67–70. [CrossRef]

Koenig, Laura L., Susanne Fuchs, Annette Gerstenberg, and Moriah Rastegar. 2020. Formant and Voice Quality Changes as a Function of Age in Women. Paper presented at the 179th Meeting of the Acoustical Society of America, Acoustics Virtually Everywhere, Online, December 7–11.

Kuznetsova, Alexandra, Per B. Brockhoff, and Rune H. B. Christensen. 2017. lmerTest package: Tests in linear mixed effects models. *Journal of Statistical Software* 82: 1–26. [CrossRef]

Lasker, Gabrial W. 1953. The age factor in bodily measurements of adult male and female Mexicans. *Human Biology* 25: 50–63. [PubMed]

Lenell, Charles, Mary J. Sandage, and Aaron M. Johnson. 2019. A tutorial of the effects of sex hormones on laryngeal senescence and neuromuscular response to exercise. *Journal of Speech, Language, and Hearing Research* 62: 602–10. [CrossRef]

Lenth, Russell. 2021. Emmeans: Estimated Marginal Means, Aka Least-Squares Means. R Package Version 1.6.2-1. Available online: https://cran.r-project.org/web/packages/emmeans/index.html (accessed on 22 November 2021).

Linville, Sue Ellen. 1992. Glottal gap configurations in two age groups of women. *Journal of Speech and Hearing Research* 35: 1209–15. [CrossRef]

Linville, Sue Ellen, and Hilda B. Fisher. 1985. Acoustic characteristics of women's voices with advancing age. *Journal of Gerontology* 40: 324–30. [CrossRef]

Linville, Sue Ellen, and Jennifer Rens. 2001. Vocal tract resonance analysis of aging voice using long-term average spectra. *Journal of Voice* 15: 323–30. [CrossRef]

Lundy, Donna S., Carlos Silva, Roy R. Casiano, F. Ling Lu, and Jun Wu Xue. 1998. Cause of hoarseness in elderly patients. *Otolaryngology—Head and Neck Surgery* 118: 481–85. [CrossRef]

MATLAB. 2018. *Version R2018b*. Natick: The MathWorks Inc.

Maxim, Jane, Karen Bryan, and Ian Thompson. 1994. *Language of the Elderly: A Clinical Perspective*. San Diego: Singular.

Michel, John F., W. S. Brown, Wojtek Chodzko-Zajko, Harry Hollien, Ruth Huntley, Joel C. Kahane, Sue Ellen Linville, and Robert L. Ringel. 1987. Aging voice: Panel 1. *Journal of Voice* 1: 53–61. [CrossRef]

Morel, Mary-Annick, and Laurent Danon-Boileau. 1998. *Grammaire de l'Intonation: L'Exemple du Français Oral*. Paris: Ophrys.

Ogden, Richard. 2001. Turn transition, creak and glottal stop in Finnish talk-in-interaction. *Journal of the International Phonetic Association* 31: 139–52. [CrossRef]

Oyer, Herbert J., and Leo V. Deal. 1985. Temporal aspects of speech and the aging process. *Folia Phoniatrica* 37: 109–12. [CrossRef]

Paulsen, Friedrich P., and Bernhard N. Tillmann. 1998. Degenerative changes in the human cricoarytenoid joint. *Archives of Otolaryngology–Head & Neck Surgery* 124: 903–6. [CrossRef]

Pontes, Paulo, Alcione Brasolotto, and Mara Behlau. 2005. Glottic characteristics and voice complaint in the elderly. *Journal of Voice* 19: 84–94. [CrossRef]

Pressman, Joel J., and George Kelemen. 1955. Physiology of the larynx. *Physiological Review* 35: 506–54. [CrossRef]

Ramig, Lorraine A., and Robert L. Ringel. 1983. Effects of physiological aging on selected acoustic characteristics of voice. *Journal of Speech and Hearing Research* 26: 22–30. [CrossRef]

Reubold, Ulrich, Jonathan Harrington, and Felicitas Kleber. 2010. Vocal aging effects on f0 and the first formant: A longitudinal analysis in adult speakers. *Speech Communication* 52: 638–51. [CrossRef]

Ringel, Robert L., and Wojtek J. Chodzko-Zajko. 1987. Vocal indices of biological age. *Journal of Voice* 1: 31–37. [CrossRef]

Rochet-Capellan, Amélie, and Susanne Fuchs. 2013. The Interplay of Linguistic Structure and Breathing in German Spontaneous Speech. Paper presented at the 14th Annual Conference of the International Speech Communication Association, Lyon, France, August 25–29; vol. 14, pp. 2014–18. Available online: https://hal.archives-ouvertes.fr/hal-00871600/ (accessed on 15 December 2021).

Russell, Alison, Lynda Penny, and Cecilia Pemberton. 1995. Speaking fundamental frequency changes over time in women: A longitudinal study. *Journal of Speech and Hearing Research* 38: 101–9. [CrossRef] [PubMed]

Schönle, Paul-Walter, and Bastian Conrad. 1985. Hesitation vowels: A motor speech respiration hypothesis. *Neuroscience Letters* 55: 293–96. [CrossRef]

Schötz, Susanne. 2006. *Perception, Analysis and Synthesis of Speaker Age. Travaux de l'Institut de Linguistique de Lund 47*. Lund: Lund University.

Schultz, Benjamin G., Sandra Rojas, Miya St. John, Elaina Kefalianos, and Adam P. Vogel. 2021. A cross-sectional study of perceptual and acoustic voice characteristics in healthy aging. *Journal of Voice*. in press. [CrossRef] [PubMed]

Scukanec, Gail P., Linda Petrosino, and Kevin Squibb. 1991. Formant frequency characteristics of children, young adult, and aged female speakers. *Perceptual and Motor Skills* 73: 203–8. [CrossRef]

Searl, Jeffrey P., Rodney M. Gabel, and J.Steven Fulks. 2002. Speech disfluency in centenarians. *Journal of Communication Disorders* 35: 383–92. [CrossRef]

Segre, Renato. 1971. Senescence of the voice. *The Eye, Ear, Nose and Throat Monthly* 50: 223–27.

Shriberg, Elizabeth. 2001. To 'errrr' is human: Ecology and acoustics of speech disfluencies. *Journal of the International Phonetic Association* 31: 153–69. [CrossRef]

Shriberg, Elizabeth E., and Robin J. Lickley. 1993. Intonation of clause-internal filled pauses. *Phonetica* 50: 172–79. [CrossRef] [PubMed]

Smorenburg, Laura, and Willemijn Heeren. 2020. The distribution of speaker information in Dutch fricatives /s/ and /x/ from telephone dialogues. *The Journal of the Acoustical Society of America* 147: 949–60. [CrossRef] [PubMed]

Sperry, Elizabeth E., and Richard J. Klich. 1992. Speech breathing in senescent and younger women during oral reading. *Journal of Speech and Hearing Research* 35: 1246–55. [CrossRef]

Stathopoulos, Elaine T., Jessica E. Huber, and Joan E. Sussman. 2011. Changes in acoustic characteristics of the voice across the life span: Measures from individuals 4–93 years of age. *Journal of Speech, Language, and Hearing Research* 54: 1011–21. [CrossRef]

Stoicheff, Margaret L. 1981. Speaking fundamental frequency characteristics of nonsmoking female adults. *Journal of Speech and Hearing Research* 24: 437–41. [CrossRef] [PubMed]

Tabain, Marija. 2003. Effects of prosodic boundary on /aC/ sequences: Acoustic results. *The Journal of the Acoustical Society of America* 113: 516–31. [CrossRef]

Taschenberger, Linda, Outi Tuomainen, and Valerie Hazan. 2019. Disfluencies in Spontaneous Speech in Easy and Adverse Communicative Situations: The Effect of Age. Paper presented at the Disfluency in Spontaneous Speech Workshop, Budapest, Hungary, September 12–14; vol. 9, pp. 55–58. [CrossRef]

Turk, L. Munir, and David A. Hogg. 1993. Age changes in the human laryngeal cartilages. *Clinical Anatomy* 6: 154–62. [CrossRef]

Vayra, Mario, and Carol A. Fowler. 1992. Declination of supralaryngeal gestures in spoken Italian. *Phonetica* 49: 48–60. [CrossRef] [PubMed]

Wagner, Anita, and Angelika Braun. 2003. Is Voice Quality Language-Dependent? Acoustic Analyses Based on Speakers of Three Different Languages. Paper presented at the 15th International Congress of Phonetic Sciences, Barcelona, Spain, August 3–9; pp. 651–54. Available online: https://www.internationalphoneticassociation.org/icphs-proceedings/ICPhS2003/papers/p15_0651.pdf (accessed on 26 July 2021).

Weirich, Melanie. 2012. The Influence of NATURE and NURTURE on Speaker-Specific Parameters in Twins' speech: Articulation, Acoustics and Perception. Ph.D. dissertation, Humboldt University of Berlin, Berlin, Germany.

Wickham, Hadley. 2016. *Ggplot2: Elegant Graphics for Data Analysis*, 2nd ed. Dordrecht: Springer International Publishing. [CrossRef]

Wolk, Lesley, Nassima B. Abdelli-Beruh, and Dianne Slavin. 2012. Habitual use of vocal fry in young adult female speakers. *Journal of Voice* 26: e111–e116. [CrossRef]

Xue, Steve An, and Dimitar Deliyski. 2001. Effects of aging on selected acoustic voice parameters: Preliminary normative data and educational implications. *Educational Gerontology* 27: 159–68. [CrossRef]

Xue, Steve An, and Grace Jianping Hao. 2003. Changes in the human vocal tract due to aging and the acoustic correlates of speech production: A pilot study. *Journal of Speech, Language, and Hearing Research* 46: 689–701. [CrossRef]

Xue, Steve An, Jack Jiang, Emily Lin, Raymond Glassenberg, and Peter B. Mueller. 1999. Age-related changes in human vocal tract configurations and the effects on speakers' vowel formant frequencies: A pilot study. *Logopedics Phoniatrics Vocology* 24: 132–37. [CrossRef]

Yamauchi, Akihito, Hisayuki Yokonishi, Hiroshi Imagawa, Ken-Ichi Sakakibara, Takaharu Nito, Niro Tayama, and Tatsuya Yamasoba. 2014. Age-and gender-related difference of vocal fold vibration and glottal configuration in normal speakers: Analysis with glottal area waveform. *Journal of Voice* 28: 525–31. [CrossRef]

Yuasa, Ikuko Patricia. 2010. Creaky voice: A new feminine voice quality for young urban-oriented upwardly mobile American women? *American Speech* 85: 315–37. [CrossRef]

Zhang, Zhaoyan. 2016a. Cause-effect relationship between vocal fold physiology and voice production in a three-dimensional phonation model. *The Journal of the Acoustical Society of America* 139: 1493–507. [CrossRef] [PubMed]

Zhang, Zhaoyan. 2016b. Respiratory laryngeal coordination in airflow conservation and reduction of respiratory effort of phonation. *Journal of Voice* 30: e7–e13. [CrossRef] [PubMed]