rijksuniversiteit
groningen
faculteit der letteren

Characterising auditory-motor adaptation of vowel production across age

Katharina Maria Polsterer

[redacted]

[redacted]

Master thesis

Research master programme in Linguistics

Track: Language and Cognition

Faculty of Arts

University of Groningen

Supervisor: Dr. Defne Abur

Second reader: Dr. Frank Tsiwah

[redacted]

Abstract

Despite the typical ease of speaking, speech production involves a complex interplay of numerous muscles. In order to speak, we move structures along the rib cage, across the larynx and up to the oral and nasal cavities in a spatially precise and temporally coordinated manner. The neurocognitive system regulating the fine motor process of speech production is called speech motor control. Sensory feedback plays an essential role in this system, allowing us to monitor our speech in real time. More precisely, we hear ourselves speak and feel the movements of the articulators. If, during speech production, errors are detected via this sensory feedback, motor commands can be corrected immediately or adapted in the long term (Guenther, 2006; Tourville & Guenther, 2011).

Speech undergoes changes throughout the lifespan. During speech and language development in childhood, the use of speech motor control mechanisms changes, typically from more towards less reliance on sensory feedback (Scheerer et al., 2016; Tourville & Guenther, 2011; Van Brenk & Terband, 2020). It is unclear, however, how the dynamics within the system of speech motor control change during adulthood. Specifically, this thesis focused on the ability to adapt stored articulatory motor programmes for speech sounds based on auditory feedback (i.e., hearing oneself speak). This learning mechanism, enabling us to learn, store, and adapt the articulation of speech sounds, is referred to as auditory-motor adaptation. Previous research has shown that the ability for immediate corrections of errors in voice pitch declines from the sixth decade of life onwards (Liu et al., 2011). This master thesis expanded on this line of research by asking whether the capacity for gradual auditory-motor adaptation of articulation changes in the course of adulthood.

To answer this question, a feedback perturbation experiment was conducted. Specifically, the first formant (F_1) of the vowel /e/ in the Italian words 'beve', 'deve' and 'vede' was gradually shifted upwards in participants' auditory feedback. The perturbation reached its maximum at an F_1 increase of 50 % relative to the respective individual's productions in baseline trials. Participants were expected to oppose the perceived error in the auditory feedback by adapting their motor commands in order to return to their baseline acoustic target.

However, the results revealed a nearly equal split in participant responses, independent of age. 41 participants opposed the F_1 increase in their subsequent productions, whereas 40 participants followed the F_1 increase. The presented data did not provide evidence of an age effect on F_1 adaptation. Thus, auditory-motor adaptation for articulation might constitute a stable mechanism throughout adulthood. However, considering previous findings, this study suggests the possibility of different aging processes for fine motor control of distinct speech components, such as voice and articulation, or different speech motor control subsystems, such as reflexive and adaptive responses to feedback perturbation. Detailed interpretations of the data, limitations of this research, and prospects for future investigations related to the current findings are thoroughly discussed.

Acknowledgments

I would like to express my heartfelt appreciation to all those whose support and contributions have been invaluable throughout the process of this thesis project:

My thesis advisor, Defne Abur, for her unwavering support and guidance throughout every step of this project, starting from Italian espresso in the mornings up to preventing this text from becoming a mess of caffeine-induced ramblings. Your enthusiasm for speech research is absolutely contagious, and makes me thrilled to continue working with you beyond this thesis. There is still so much to explore!

To the remarkable group of people in the Speech Lab Groningen who have affiliated me from the beginning of my master studies. This thesis is based on a collaborative effort, and an adventure that I did not image I would experience when I joined the lab: travelling to a festival in Italy together. I am deeply grateful to be a part of this group. I would like to specifically point out two people whose support has been extraordinarily valuable over the past years as well as on the path to completing this thesis: Martijn Wieling and Teja Rebernik.

I also want to thank Frank Tsiwah for his openness to help out, whether it was taking the time to conduct mock interviews or reading and assessing this thesis.

No acknowledgments section would be complete without mentioning the solid support of my partner and family. Thank you, Michi, for being on this rollercoaster together with me. I owe you much more than a spot in this section. And thank you to my family, who is enduring the ideas I set my mind on. You never fail to remind me that there is a world beyond academia, where I can always come back home to, calm down in the vineyards, and go Stamm obutzn. Do kau ma mocha wos ma wü, do bi i her, do kea i hi.

Table of contents

1. Introduction 1
2. Theoretical background 3
    2.1. Filtering sounds: The articulatory system 4
    2.2. Sensorimotor control of the articulatory system 6
        2.2.1. The DIVA model: Directions into Velocities of Articulators 7
        2.2.2. Fine motor control via auditory information: Auditory-motor control 9
        2.2.3. Changes of auditory-motor control across age 10
    2.3. The present study 16
3. Methods 18
    3.1. Participants 18
    3.2. Data collection 20
        3.2.1. Pre-experimental procedure 20
        3.2.2. Experimental setup 20
    3.3. Data analysis 22
        3.3.1. Pre-processing 22
        3.3.2. Statistical analysis 23
4. Results 24
    4.1. Exploratory analysis 25
    4.2. Hypothesis testing 29
5. Discussion 34
    5.1. Directions of adaptive responses: Following vs. opposing the perturbation 35
        5.1.1. Producing the Italian /e/: Allophonic variability and phoneme boundaries . 36
        5.1.2. The structure of the experimental stimuli 39
        5.1.3. Summary of potential confounds 40
    5.2. Aging 40
        5.2.1. Variability of (adapted) motor programmes 44
        5.2.2. Aging and Sex 45

5.2.3. Implications for the DIVA model 46

5.3. Limitations and avenues for future research 46

6. Conclusion 49

7. References 50

8. Appendices 59

8.1. Appendix 1: Individual mean baseline F_1 59

8.2. Appendix 2: Response magnitudes (% F_1 deviation in the hold phase) 60

1. Introduction

Speech is a fundamental aspect of human communication, relying on the intricate interplay between our ability to produce and perceive spoken language. The coordination between our fine motor and auditory systems allows us to adapt articulatory movements based on their acoustic results. Early on, children use this mechanism to acquire speech, a process which is apparent as they babble. In that developmental phase, children pick up speech sounds in their environment, and try to imitate those sounds. Consequently, they hear their own speech, which they compare to the target sounds. Through multiple repetitions, they are able to adapt their articulatory movements step by step and eventually come closer to their auditory targets. Articulatory movement patterns corresponding to desired speech sounds are then stored and recurrently practiced. This learning mechanism is referred to as auditory-motor adaptation (Guenther, 2006; Tourville & Guenther, 2011).

As we grow older, more specifically beyond childhood, the auditory-motor adaptation mechanism is not only preserved but remains in use. For instance, we can learn the sounds of other languages, and imitate other accents or dialects. In everyday conversations such as telephone calls or in noisy environments, we may need to temporarily adjust our speech in order to be intelligible. However, age-related changes in this control mechanism may affect our ability to adapt and refine speech movements (Li et al., 2018). In light of our aging society (Eurostat, 2023)[1], understanding the impact of aging on speech is crucial in obtaining a comprehensive picture of the human aging process and its implications for quality of life including our ability to speak.

However, the exact way in which aging affects the control of speech movements through sensory information remains yet unclear. The present master thesis therefore explored one specific aspect of the human aging process: that is, the way we use the perceptual system to flexibly adapt stored motor commands for speech articulation. To this

---

[1] On 1 January 2022, people aged 65 or older constituted 21.1 % of the EU's population, which is an increase of 3.1 % compared to the year 2012 (Eurostat, 2023).

end, the capacity for vowel adaptation specifically based on the auditory perception route was investigated experimentally across adulthood.

2. Theoretical background

Human speech is generated through a complex interplay between three components: the respiratory system, the laryngeal system, and the articulatory system. Speech is produced as air flowing through these systems is modulated. Both vowels and consonants are produced via that pathway and in that manner. The current thesis specifically focuses on the production of vowels.

The respiratory system comprises the lungs, as well as the skeletal and muscular structures that support breathing. When speaking, expiration is prolonged compared to tidal breathing, which allows for the regulation of air pressure that reaches the laryngeal system (McFarland, 2015). According to the power-source-filter model of speech production, this system provides the 'power' for speaking (Titze, 1994).

Air from the respiratory system passes through the laryngeal system. More specifically, the laryngeal system refers to the larynx including the vocal folds (McFarland, 2015). As air flows from the lungs through the trachea, it passes the space between the vocal folds, which is termed the glottis. For tidal breathing, the vocal folds are abducted, meaning that they are wide apart providing an open glottis. When generating voice, on the other hand, the vocal folds are adducted. That is, the glottis is closed, so that the air exerted by the lungs is temporarily stopped from flowing. As soon as subglottal air pressure exceeds the air pressure above the vocal folds, the flexible covering layers of the vocal folds are pushed apart. Due to the resulting change in air pressure states, the vocal folds are drawn back together by Bernoulli's principle along with an elastic force. As this process repeats rapidly, the vocal folds vibrate and cause air pressure waves in the air flow coming from the lungs (Jiang et al., 2000). The number of vibrations per second constitutes the fundamental frequency of a person's voice (f_o). The laryngeal system is therefore referred to as the 'source' for speech production (Titze, 1994).

Sound coming from the laryngeal system (voice) is modified as it travels through the articulatory system (McFarland, 2015). In particular, the voice signal gets amplified at specific locations along the vocal tract depending on its shape. In that way, components of the source signal are shaped to produce different vowels. The power-source-filter theory of speech therefore terms the vocal tract the ‘acoustic filter’ for speech production (Titze, 1994). This filter is the main focus of the present thesis.

2.1. Filtering sounds: The articulatory system

The vocal tract consists of the pharyngeal cavity, the nasal cavity, and the oral cavity (McFarland, 2015). Depending on the configuration of the vocal tract, some frequency components of the source signal are enhanced, while others are dampened (Tokuda, 2021). The configuration is changed through displacements of the articulators, which include the tongue, the jaw, and the lips (McFarland, 2015; Tokuda, 2021). These displacements create different locations of constrictions, so that sound pressure changes as air flows through the vocal tract. This results in resonant frequencies along the vocal tract, where sounds are amplified, that create so-called ‘formant frequencies’ in the resulting speech signal (Tokuda, 2021). Therefore, acoustic formant frequencies in the speech signal contain important information about the shape of the vocal tract.

Formant frequencies, or formants, crucially contribute to vowel identity, and thus perception by the listener as they shape the acoustic quality of different vowels. Each vowel phoneme is characterized by several formant values typical for a specific vowel in a speaker’s language. These different formants are modulated at distinct parts of the vocal tract. The first formant (F_1) provides especially critical information to the listener. It primarily reflects the volume of the pharyngeal and oral cavity, modified by a constriction between the tongue and the palate. Higher F_1 frequencies, such as in /a/ as opposed to /i/, are produced with the tongue in a relatively low position within the oral cavity. Lower F_1 frequencies, such as in /i/ as opposed to /a/, on the other hand, are produced with the tongue positioned high, and thus close to the palate. Hence, a larger physiological space in the vertical dimension of the vocal tract is related to lower F_1 frequencies. On the contrary, the second formant (F_2) is inversely correlated to the horizontal dimension of the oral cavity. That is, a more frontal position of the tongue causes a higher F_2 frequency. Accordingly, the vowel /i/ is characterized by a higher F_2 than the vowel /u/ since for the latter, the tongue is relatively more retracted (Liberman et al., 1967). While there are more than these two formants, mostly F_1 and F_2 are discussed in the literature on speech motor control (Cai et al., 2023; Daliri & Dittman, 2019; Houde & Jordan, 1998; Lametti et al., 2018) as they are the most dominant in the acoustic signal. This thesis focuses on F_1 perturbation and adaptation. As F_1 corresponds to the height of the tongue position, participants are expected to adapt this specific parameter in response to acoustic shifting of the auditory feedback. Specific formant values for distinct vowel phonemes vary per language. The average F_1 of the standard Italian /e/ has been reported to be between 390 and 439 Hz, depending on the preceding consonant (Esposito, 2002). In the experiment conducted for this thesis, participants produced words containing /e/. They perceived the F_1 frequency in their own speech higher than they would have expected based on their tongue height.

In the present thesis, speech adaptation responses to this perturbation were studied across age, while potential differences by sex were explored. When comparing individuals across age and sex, structural and functional differences relevant to speech production need to be considered. As humans age, vocal tract structures change gradually, which affects acoustic characteristics of speech. When comparing 38 young ($\bar{x}$ male = 21.89 years of age; $\bar{x}$ female = 22.21 years of age) to 38 older ($\bar{x}$ male = 71.32 years of age; $\bar{x}$ female = 74.21 years of age) adults using acoustic reflection technology, Xue and Hao (2003) found that the oral cavity lengthens with age, whereas no differences were observed in the pharyngeal cavity. This anatomical change was correlated with a lowering of the F_1 in both male and female participants. Later work demonstrated evidence in support of this (Watson & Munson, 2007) as well as conflicting findings (Traub-Eichhorn et al., 2018). Watson and Munson (2007) additionally found F_2 to be lower in older ($\bar{x}$ = 76 years of age) than younger ($\bar{x}$ = 23.3 years of age) adults leading to a group difference in the shapes of the acoustic vowel space. On the contrary, Traub-Eichhorn et al. (2018) investigated age-related changes in F_1 to F_4 of 43 male and 53 female participants (age range = 20 – 92 years) but did not confirm systematic age-related changes of the acoustic vowel space as suggested in earlier studies. Besides inherent dimensional changes of the vocal tract, movement constraints originating from the temporomandibular joints might cause vowel space alterations, and especially vowel height (i.e., F_1) alterations, in older people. However, the typical aging process does not appear to restrain the function of healthy temporomandibular joints in 65- and 75-year-old adults (Unell et al., 2012).

Tongue muscle mass and thus strength decrease in older age (Lamster et al., 2016). Watson and Munson (2007) consider this factor irrelevant for vowel production stating that only a low level of muscle strength is required for speech production. Braun and Friebis (2009), on the other hand, argue that a decrease of muscle mass decelerates articulatory movements of the tongue and leads to imprecision in older individuals' speech. Empirical research tested whether age-related decline in muscular strength of the lips and tongue, muscular endurance, or tactile sensitivity affects speech rate or accuracy. The results revealed that only decreased lip endurance impacts the accuracy of speech (Bilodeau-Mercure & Tremblay, 2016). Overall, older adults have been found to show an increased error rate in speech as opposed to younger adults (Tremblay et al., 2018), which by inference cannot be fully traced back to physiological changes of the vocal tract.

Structural and functional aging of the vocal tract as outlined in this section do not seem to differ considerably between males and females. Specifically, the dimensions of the vocal tract (Xue & Hao, 2003) as well as the acoustic vowel space (Traub-Eichhorn et al., 2018) develop similarly in both sexes. Therefore, potential sex-related differences in the aging process of vowel adaptation are likely not due to the discussed factors.

2.2. Sensorimotor control of the articulatory system

In order to produce fluent speech, the muscles and structures involved in articulation (i.e., the filter) as well as respiration and voice production (i.e., the source) need to be controlled and coordinated in a precisely timed manner. For articulation, sequences of muscle activations and deactivations initiated by the central nervous system cause the tongue, the lips, the jaw, and the velum to move in a multidimensional space. Typically, the distal end of each articulator, such as the tongue tip, is precisely positioned. Consequently, the relation between possible articulator configurations and the produced acoustic signal is of a non-linear nature (Parrell, Lammert, et al., 2019).

A range of contemporary neurocomputational models attempt to explain this intricate mechanism (see review by Parrell, Lammert, et al. (2019) for a comparative overview). Widely-used models include *Directions into Velocities of Articulators* (DIVA; Guenther, 2006; Guenther & Vladusich, 2012; Tourville & Guenther, 2011), *State Feedback Control* (SFC; Houde & Nagarajan, 2011), and *Feedback Aware Control of Tasks in Speech* (FACTS; Parrell, Ramanarayanan, et al., 2019), amongst others. Different models of speech motor control coincide in the underlying mechanistic principles of motor control. That is, desired verbal output is prepared through higher-level linguistic processes. Lower-level speech production comprises a *planner* generating a speech target based on the linguistic plan, a *controller* generating a corresponding motor programme, and a *plant* executing those motor programmes and hence moving into different states (Parrell, Lammert, et al., 2019). For theoretical background and placement within exiting literature, this thesis is specifically based on the neurocomputational model *Directions into Velocities of Articulators* (DIVA; Guenther, 2006; Guenther & Vladusich, 2012; Kearney & Guenther, 2019; Tourville & Guenther, 2011).

2.2.1. The DIVA model: Directions into Velocities of Articulators

The DIVA model assumes a close interplay of sensory information and fine motor skills to correct speech movements instantly but also adapt movement programmes in the long term. When producing speech, abstract linguistic units activate associated neurons in the speech sound map. In more theoretical terms, this process constitutes the transformation of a phonological to a phonetic representation. While linguistic units can range in size from phonemes to short sequences of syllables, the DIVA model considers syllables as the typical representations on the speech sound map (Guenther & Vladusich, 2012). In a similar vein, Levelt (1989) assumes a mental syllabary at the intersection of phonological and phonetic processes in his influential theory of speech production. A line of empirical research has provided evidence for such a neurocognitive storage of syllables (e.g., Cholin & Levelt, 2009). Accordingly, Karlin et al. (2021) demonstrated that representations of temporal patterns go beyond single phonemic segments. The authors employed a feedback perturbation paradigm, during which they manipulated consonant durations in participants' auditory feedback (i.e., what participants heard themselves say in near real time) as they were pronouncing noun phrases such as 'a capper'. Their participants tended to adapt proportionate durations across their speech productions rather than absolute vowel lengths in response to the perturbations. It was concluded that the speech sound map encodes acoustic characteristics beyond single phonemes.

Neuroimaging studies have located the speech sound map in the left ventral premotor cortex (Guenther, 2016; Tremblay et al., 2017) as well as the left inferior frontal cortex (Guenther, 2016; Markiewicz & Bohland, 2016). Both areas thus pertain to the neural speech motor control network, which overall consists of cortical alongside subcortical structures and regulates targeted movements of the articulators (Guenther & Vladusich, 2012). The speech sound map provides input to two subsystems of the speech motor control network, which work in tandem: feedforward and feedback control. Stored speech sounds include information on feedforward motor commands as well as speech targets to which sensory feedback is compared (Tourville & Guenther, 2011).

Feedforward control. Activated neurons of the speech sound map send time-varying feedforward motor commands to the articulator velocity and position maps located bilaterally in the ventral motor cortex and cerebellum. The latter maps represent current configurations of the articulators (Guenther & Vladusich, 2012; Tourville & Guenther, 2011). As speech musculature is activated by those feedforward commands, speech is produced. While evidence suggests that the speech sound map accounts for syllables, motor commands for individual phonemes are malleable. Adaptations of motor programmes for individual phonemes are not restricted to specific syllables on which the adaptation was trained but can be generalized to other contexts and acoustically neighbouring phonemes (e.g., /ɛ/ in (Houde & Jordan, 1998), /iau/ in (Cai et al., 2010)).

Feedback control. During speaking, acoustic as well as somatosensory signals are perceived by the speaker and registered in sensory state maps. Auditory feedback information is sent back to the auditory cortex, spanning Heschl's gyrus and the posterior superior temporal gyrus (pSTG), via the cochlea and the VIIIth cranial nerve (Guenther, 2016; Hickok & Poeppel, 2004). By means of magnetoencephalography, the auditory cortex was shown to differentiate externally and internally generated speech through the magnitude of activation (Heinks-Maldonado et al., 2006). Accordingly, Daliri and Dittman (2019) reported that behavioural error sensitivity is higher for errors attributed to speakers' own actions compared to externally induced errors. At the same time, tactile information is received by sensory receptors of the articulators and sent to the primary somatosensory cortex via the Vth, IXth, and Xth cranial nerve (Guenther, 2016; Hickok & Poeppel, 2004). Auditory and somatosensory feedback are compared to expectations based on the initiated feedforward motor commands. This comparison allows for spontaneous corrections of produced speech. That is, the feedback control loop is able to generate immediate corrective motor commands, if required (Guenther, 2006).

The different components of DIVA outlined here can be studied experimentally using behavioural besides neuroimaging measures. More specifically, gradual or sudden manipulation of sensory feedback typically triggers compensatory modifications of articulatory movements (Burnett et al., 1997; Cai et al., 2008; Houde & Jordan, 1998; Villacorta et al., 2007). The experiment presented in this thesis employed such a behavioural experiment to test adaptations of feedforward motor control, based specifically on auditory feedback.

2.2.2. Fine motor control via auditory information: Auditory-motor control

Speech motor control generally includes the ability to generate spontaneous motor corrections for speech errors, as well as long-term adaptations of stored motor commands. Both types of motor compensation are based on continuous integration of sensory feedback information in the central nervous system. Thereby, incongruencies between the expected and the perceived sensory feedback are recognized and transformed to motor equivalents. Auditory-motor control specifically refers to compensatory mechanisms that draw on errors detected in the auditory feedback.

Auditory error maps. Both feedback control and adaptation of feedforward control rely on the detection of errors in the auditory feedback. Based on activations in the speech sound map, associated neurons in the auditory target map are activated, which represent expectations for perceived feedback information (i.e., a ‘forward model’). While the target map inhibits auditory error maps, input from the auditory state map excites this error map (Tourville & Guenther, 2011). Auditory target and state maps have been suggested to lie in close proximity to the auditory error map in Heschl’s gyrus and the pSTG (Guenther, 2016). The difference between feedforward-based expectations and feedback-based information is represented by activation in the auditory error map (Heinks-Maldonado et al., 2006).

As auditory errors for vowels reflect errors in formant frequencies on the vowel space, they can be quantified using behavioural experiments. The behavioural ability to discriminate auditory stimuli is referred to as ‘auditory acuity’. Higher auditory acuity is typically interpreted as more precise representations on the speech sound map. Perkell et al. (2004) demonstrated that auditory acuity varies between participants and correlates with productive acuity. That is, the researchers recorded acoustic and articulatory signals of young adults pronouncing minimal pairs (CVC structure) contrasting /ɑ/ and /ʌ/, alongside /u/ and /ʊ/. Participants who showed high auditory acuity, also showed low productive variability within tokens as well as large articulatory and acoustic differences between phonemes. This result could be replicated later on (Franken et al., 2017). These findings confirm the relation of speech sound representations for speech production and perception.

Auditory errors mapped to motor changes. When an auditory signal is perceived, auditory cortex activity is passed on to the speech motor control system (Tourville & Guenther, 2011). If an auditory error is detected, the feedback control mechanism is able to initiate an instantaneous corrective motor command. That is, the auditory error is transformed to a corrective, also referred to as reflexive, motor sequence. The latter is projected to the articulator velocity and position maps to be executed (Guenther & Vladusich, 2012; Tourville & Guenther, 2011). The magnitude of reflexive corrections to F_1 feedback perturbations was found to positively correlate with auditory acuity for F_1 distinctions (Lester-Smith et al., 2020).

Motor changes incorporated into the feedforward system. If auditory errors occur repeatedly and consistently (Franken et al., 2019), corresponding motor changes are incorporated into the speech sound representations in the long term (Guenther, 2006). That is, feedforward motor commands can be adapted. The ability for motor adaptation can be investigated by experimentally shifting formants in the perceived auditory feedback. The relation between auditory acuity and motor adaptation, however, is unclear. Smaller perceptual targets (Daliri & Dittman, 2019) as well as higher auditory acuity (Villacorta et al., 2007) have been reported to correlate with greater motor adaptation in response to formant shifting. Conversely, Lester-Smith et al. (2020) found no relation between auditory acuity to differences in F_1 and adaptation to upward shifting of F_1. On the other hand, Lester-Smith et al. (2020) identified a correlation between reflexive responses to F_1 perturbations and auditory-motor adaptation of vowels. That is, larger reflexive responses were associated with larger adaptive responses. This result is interpreted with caution by Lester-Smith and colleagues as previous findings by Franken et al. (2019) do not coincide with this outcome. The latter authors concluded that reflexive compensations and gradual adaption might constitute two distinct mechanisms.

2.2.3. Changes of auditory-motor control across age

Auditory-motor control mechanisms regulate speech production throughout the lifespan (e.g., Hu et al., 2023; Van Brenk & Terband, 2020). However, there seems to be a change in how they are used. While children primarily rely on feedback control during early learning phases, young adults generally rely more strongly on feedforward control as speech motor programmes have been stabilized (Guenther, 2006; Tourville & Guenther, 2011).

Development in childhood. Speech motor programmes are learnt during the infant babbling phase. Children imitate auditory speech input from their environment to develop corresponding motor actions through multiple repetitions (Hickok & Poeppel, 2004; Tourville & Guenther, 2011). As a consequence of each speech production, they perceive auditory feedback, which is processed in the sensory state map of the STG (Guenther, 2016; Hickok & Poeppel, 2004). Feedback integration via comparison to the forward model allows the child to adapt their motor command for the next repetition. Thereby, children gradually learn to produce target sounds as they practice through babbling (Guenther, 2006). They store motor programmes pertaining to desired acoustic outputs including accepted variability in the speech sound map. However, variability in production is larger in children (5 to 10 years of age) compared to adults (their mothers) in temporal as well as spatial domains of articulation (Koenig et al., 2008) suggesting ongoing plasticity of their speech sound maps. Van Brenk and Terband (2020) compared 4- to 9-year-old children to 18- to 29-year-old adults in their ability to adapt their vowel productions within CVC words in response to shifts in the F_1 and F_2. Similar to Koenig et al. (2008), they found that, while children overall produce vowels more variably, they adapt more to the experimental manipulation. Thus, the authors conclude that variability is associated with children's higher capacities for auditory-motor plasticity and learning.

Auditory-motor control in adulthood. Yet in adulthood, established and stabilized motor programmes still remain adaptable via the sensory-motor route (Cai et al., 2010; Houde & Jordan, 1998). Motor commands can be adapted for specific phonemes (Houde & Jordan, 1998) but also for time-varying linguistic units (e.g., triphthongs) (Cai et al., 2010). While adaptations of motor commands for specific vowels can be generalized to other vowels, such generalization depends on the acoustic distance and kinematic differences between the manipulated vowel and other vowels (Cai et al., 2010).

Adults are able to use auditory-motor control not only to adapt stored motor programmes but also to learn new speech sounds. As exemplified by Hickok and Poeppel (2004), adults' ability to acquire and adapt motor commands for speech is evident by their ability to build new articulatory sequences for pseudowords. When learning a new language during adulthood, the speech sound map's repertoire may need to be expanded using a similar auditory-motor learning mechanism as during childhood (Cai et al., 2023). Cai et al. (2023) studied first language (L1) speakers of Chinese, who had acquired English as a second language (L2). They compared behavioural besides neurophysiological responses to suddenly shifted feedback for voice pitch control in participants' L1 and L2. As voice pitch fulfils different linguistic functions in the two involved languages, L2 learners were required to complement their neurocognitive speech sound representations to account for those different functions. The study discovered diverging neural processing for L1 and L2 and similar, yet less efficient, reflexive speech error correction in the L2.

Reflexive and adaptive auditory-motor control seem to vary considerably between adults. Researchers have looked at a range of potential factors to explain this variability. For instance, personal weighing of sensory (i.e., auditory and somatosensory) feedback might contribute to inter-speaker variability of speech motor adaptation in response to auditory perturbation. Lametti et al. (2012) tested 75 native English speakers, aged between 18 and 40 years, on their productions of /ɛ/ versus /æ/. By employing somatosensory alterations of the jaw as well as downward shifting of the F_1, they found some participants relied mostly on auditory feedback, while others favoured somatosensory feedback and others didn't show preferences. Based on a later study, it was proposed that involvement of somatosensory and auditory feedback might additionally depend on the intended vowel itself. That is, English speaking participants produced all vowel phonemes of English in the same phonological context: /hVd/. Shifting the F_1 in upward or downward direction yielded most pronounced responses for open vowels. Adaptation was smaller for closed back and the smallest for closed front vowels. The authors interpreted this finding such that somatosensory feedback may be most involved in closed front vowels (Mitsuya et al., 2015). This explanation corresponds to general principles of motor control, which assume that the distal end of articulators (i.e., the *plant*) constitute the kinematic focus (Parrell, Lammert, et al., 2019), suggesting high acuity and precision at the tongue tip.

Beyond inter-speaker variability, adaptation to auditory perturbations can be influenced by methodological choices of the researchers. This includes the direction of feedback shifting (e.g., increase in F_1 as done by (Abur et al., 2021) or decrease in F_1 as done by (Cai et al., 2010)), as well as the distance between the manipulated vowel and the neighbouring vowel in shifting direction (Kothare et al., 2020). Accordingly, in everyday life, articulatory adaptation likely also depends on factors such as acoustic distances of existing representations are speech sounds to be learned.

Aging processes. While variability in speech production has been attributed to many factors, aging processes might influence auditory-motor control in adults. However, aging processes of auditory-motor control are embedded within a broader multifaceted and dynamic developmental process that takes place throughout the lifespan. Aging brings about a multitude of changes in bodily structures and functions that can influence speech production as well as perception in various ways. Therefore, the bigger picture of aging needs to be kept in mind when examining the specific processes for auditory-motor control. Specifically, degradation of anatomical structures can have noticeable influences on speech. Muscles across the body degrade in mass along with strength from approximately the fourth decade of life onwards (Nair, 2005). Accordingly, orofacial muscle endurance decreases, which may influence the ability to produce clear and distinct speech sounds (Bilodeau-Mercure & Tremblay, 2016). Moreover, sensory functions relevant for speech may similarly decay with increasing age. Both somatosensory (Ketcham & Stelmach, 2004) and auditory (Slade et al., 2020; Tremblay et al., 2003) capabilities frequently decrease, which may lead to variability in reliance on one or the other sensory domain. As hair cells in the peripheral auditory system degenerate, the sensory function of hearing tends to gradually decline, with higher frequencies being affected more strongly than lower frequencies. While adults at 20 years of age typically hear 8,000 Hz tones at an amplitude of 0 dB hearing level (HL), adults at 70 years of age typically hear that same frequency with 50 (male) or 40 (female) dB HL (Slade et al., 2020). Furthermore, the auditory cortex seems to process information more slowly in older (61 – 79 years of age) as compared to younger (19 – 32 years of age) adults (Tremblay et al., 2003). Hearing loss additionally leads to a decrease in the amplitude of N1 responses to speech stimuli (Tremblay et al., 2003). This event-related potential (ERP) has been associated with the detection of feedback errors in previous literature (Behroozmand et al., 2011; Li et al., 2018). As a result, speech discrimination and consequently speech-in-noise perception may be hampered in older individuals (Slade et al., 2020). As accurate speech perception is crucial for the auditory-motor adaptation process, older adults (63 – 74 years of age) with hearing loss at frequencies relevant for speech (250 to 4,000 Hz) show reduced recruitment of articulatory motor areas during speech processing (Panouilleres & Mottonen, 2018). These closely interacting mechanisms covering sensory and motor systems shape the human development, and importantly speech production across age.

Specific aging processes of speech can be observed behaviourally both from a kinematic and an acoustic perspective. Physiological changes to relevant orofacial functions in relation to aging have been documented. For example, speech movements of the tongue show slower peak velocities and more asymmetric kinematic patterns in older (70 – 80 years of age) compared to younger (20 – 30 years of age) adults. That is, movement deceleration is prolonged, whereas acceleration is shortened (Hermes et al., 2018; Mücke et al., 2020) coinciding with aging patterns in general motor control (Ketcham & Stelmach, 2004). Prolonged deceleration of the tongue before reaching the articulatory target position is likely due to increased time requirements for the incorporation of auditory feedback (Hermes et al., 2018). In accordance with this explanation, aging has been demonstrated to lead to delayed auditory processing of speech sounds (Matilainen et al., 2010). Longer and more variable movement times in older (mean age = 68 years) compared to younger (mean age = 26.8 years) adults are further influenced by structural complexity of the sequence to be produced. This difference between age groups was associated with an expansion of recruited neural resources to the right posterior cingulate cortex (Tremblay et al., 2017). As this area is not considered as being typically involved in speech motor control (Guenther, 2016) but the default mode network (Buckner et al., 2008), its activation may constitute an age-related compensatory mechanism (Tremblay et al., 2017). In line with the observed kinematic changes, older age was found to be correlated with acoustically slower and more variable speech in 18 to 83 year-old persons (Tremblay et al., 2018). This finding concurred with previous research comparing younger (18 – 39 years of age) and older (66 – 85 years of age) groups of speakers (Bilodeau-Mercure & Tremblay, 2016). Acoustic accuracy of speech production seems to be influenced by the phonetic manner of target sounds. Specifically, nasality has shown age-related decline in accuracy (Bilodeau-Mercure & Tremblay, 2016). Nasal sounds are productively complex as they require precise coordination of the velum, lips and tongue. The vulnerability of nasal sounds may also be related to an age-related loss of structural elasticity in the oral cavity entailing reduced acoustic resonance (Braun & Friebis, 2009). Thus, productive complexity as well as anatomical degradation may change produced speech acoustically in older age. Despite these general patterns seen in aging speech, between-speaker variability in speech production seems to increase as a function of age (Hermes et al., 2018). This increase may be related to inter-speaker differences in sensory and motor degradation.

While diverse literature on age-related changes relevant to speech is available, little is yet known about the effect of aging on neural motor control of speech (Hu et al., 2023). However, previous research by Tremblay et al. (2018) suggests that speech planning and execution processes are affected by advancing age. More precisely, the authors demonstrated that the production of non-words, which requires online planning of speech motor programmes, is performed with increasing error rates with advancing age (18 – 83 years of age). Moreover, speakers tended to slow down their articulation during online speech planning, a strategy which tended to gradually decline at older age. In a similar vein, feedback control as described by the DIVA model seems to be subject to aging. In previous research, 19- to 75-year-old speakers of Mandarin Chinese produced the vowel /u/, and perceived sudden upward shifts in their fundamental frequency (f_o). The magnitudes of the participants' reflexive compensations were reported to increase with age up to 51 years, after which they reached a plateau. On the contrary, magnitudes of reflexive compensations decrease again from age 61 onwards. The maximum values of response magnitudes correlated with higher sensitivity to the direction and magnitude of f_o shifting (Liu et al., 2011). This age effect was reflected by shorter P2 latencies in younger (aged 19 – 25) compared to older (aged 60 – 73) adults (Li et al., 2018). The P2 potential is assumed to show auditory-motor integration processes (Behroozmand et al., 2011; Li et al., 2018). Thus, aging seems to increase the time needed to incorporate auditory information and transform detected errors to reflexive motor corrections. This proposal is in line with behavioural findings by Hermes et al. (2018) and Mücke et al. (2020), who suggested that those increased neural computations are evident in prolonged deceleration phases during articulation.

Aging processes of auditory-motor control also seem to differ between sexes. Li et al. (2018) reported that P2 amplitudes in response to f_o upward shifting were significantly smaller in older males as well as young females compared to young males. Additionally, N1 amplitudes were smaller in young female than young male participants, while females generally showed decreased latencies in N1 responses compared to males (Li et al., 2018). These results suggest sex-dependent auditory-motor processing at a young age, followed by distinct aging trajectories per sex.

In sum, Li et al. (2018) as well as (Liu et al., 2011) found age-related changes in f_o control in native speakers of Mandarin Chinese. f_o, or voice pitch, has an additional function in Mandarin Chinese: Lexical tones are used to distinguish meanings. Thus, for the speaker in the two above-mentioned studies, f_o fulfils a function similar to formants in addition to conveying social information. Accordingly, others found articulatory differences between age groups (Hermes et al., 2018; Mücke et al., 2020) and alterations in timing and accuracy with age (Bilodeau-Mercure & Tremblay, 2016; Tremblay et al., 2018; Tremblay et al., 2017).

2.3. The present study

Taken together, recent scientific findings indicate a relation between typical aging and variations in auditory-motor control, specifically reflexive feedback control of voice production. The influence of aging on auditory-motor adaptation, however, has not been clarified thus far. Moreover, evidence suggests age-related changes in articulatory patterns, but less is known about the influence of age on auditory-motor control for articulation. The currents master thesis addressed these two gaps to add further evidence of the aging processes of auditory-motor control of speech. Specifically, the presented study aimed at investigating auditory-motor adaptation for articulation through answering the following research question:

How does typical aging affect auditory-motor adaptation in the articulatory domain?

Based on the literature reviewed above (Li et al., 2018; Liu et al., 2011; Tremblay et al., 2018; Tremblay et al., 2017), three hypotheses were put forward. (1) It was hypothesized that, overall, speakers flexibly adapt articulatory motor commands by opposing errors detected in auditory feedback as observed in prior research studying various populations (e.g., Abur et al., 2021; Houde & Jordan, 1998; Lester-Smith et al., 2020). (2) Furthermore, it was hypothesized that increased age presents with declined auditory-motor adaptation due to reduced cortical integration of sensory feedback or increased reliance on feedback as opposed to feedforward control as suggested by theoretical models of speech motor control (Guenther, 2006; Houde & Nagarajan, 2011). (3) In addition, based on previous findings (Li et al., 2018), this aging effect was hypothesized to differ between sexes in early adulthood, while such differences may even out with higher age.

Hypotheses on auditory-motor adaptation for articulation can be tested using an adequate variety of the auditory feedback perturbation paradigm, which has frequently been adjusted to specific research questions in previous work. That is, auditory-motor adaptation has been studied by upward and downward shifting of F_1 and F_2 frequencies in vowels (Houde & Jordan, 1998), upward and downward shifting of F_1 frequencies in triphthongs (Cai et al., 2010), upward shifting of only F_1 frequencies in vowels (Abur et al., 2021; Lester-Smith et al., 2020), predefined (i.e., formant-clamp) versus online formant shifting (Daliri & Dittman, 2019), temporal manipulations of phoneme durations (Karlin et al., 2021; Oschkinat & Hoole, 2020), and kinematic disruptions of jaw movements (Nasir & Ostry, 2009), amongst others. To test the hypotheses stated above, gradual upward shifting of F_1 in the vowel /e/ was employed. This approach was chosen as previous research has shown that perturbation of one formant is sufficient to initiate auditory-motor adaptation. Moreover, a phonological category switch due to the increased F_1 would have resulted in meaningful words. This perturbation was conducted in four phases: Baseline, ramp, hold, and after-effect.

Assuming the above hypotheses are true, the following outcomes were predicted for the present study. (1) In response to gradual F_1 increase, it was predicted that participants would gradually decrease F_1 productions, with the highest deviance from the individual baseline F_1 for /e/ seen in the hold phase. This pattern is predicted by the DIVA model (Guenther, 2006; Tourville & Guenther, 2011) and empirically confirmed by previous research (Houde & Jordan, 1998; Villacorta et al., 2007). (2) With respect to aging, it was predicted that articulatory adaptation (i.e., the deviance between F_1 production in the hold compared to the baseline phase) would increase in magnitude up to the fifth decade of life, and decrease from the sixth decade of life onwards. This prediction was based on the auditory-motor aging pattern delineated by Liu et al. (2011). (3) In terms of differences between males and females, based on Li et al. (2018), it was predicted that age-related changes to F_1 adaptation (i.e., magnitude of the deviance in the hold compared to baseline phase) would differ between sexes.

3. Methods

The data for this study was collected at the Festival della Scienza 2022, which took place from 20th October until 1st November 2022 in Genoa, Italy. This festival constitutes a platform for scientific outreach to the wider non-expert public, and therefore allowed us to recruit participants across a broad age range. A mobile laboratory van holding a booth, dampened to -40 dB, was used to conduct the experiment at the festival. All study procedures were approved by the Research Ethics Review Committee (CETO) of the Faculty of Arts of the University of Groningen (reference number 82182577) and complies with the Declaration of Helsinki.

3.1. Participants

Throughout the festival, 93 individuals aged 16 to 82 years ($\bar{x}$ = 39.06, *sd* = 17) were recruited. To be included in the current study, speakers needed to be native Italian speakers and have no history of speech, language, hearing, or neurological disorder. The first criterion was chosen because previous research has indicated differences in speech motor control between first languages and second languages acquired late (Cai et al., 2023). The second inclusion criterion relates to the known influences of speech, language, hearing, and neurological disorders on speech motor control (Abur et al., 2021; Ranasinghe et al., 2017; Terband et al., 2014). In accordance with these criteria, six individuals were excluded due to a history of speech disorders. Two individuals were excluded since Italian was their second language. Furthermore, there was one participant who indicated 'other' as their sex. Since sex was included as a variable in the statistical analysis, this person needed to be excluded given that there was not enough data on their group. In the final sample, 81 participants were included for the current study. Figure 1 shows the age and sex distribution among the participants.

Demographic data corresponding to our final participant sample are summarized in Table 1. Three participants indicated to be bilingual native speakers of Italian. Their other native languages were Albanian, English, and French. Furthermore, 67 participants had learned second languages later in life. The most common among those languages were English (*n* = 59), French (*n* = 21), Spanish (*n* = 17), and German (*n* = 11).

Figure 1

Age distribution by sex

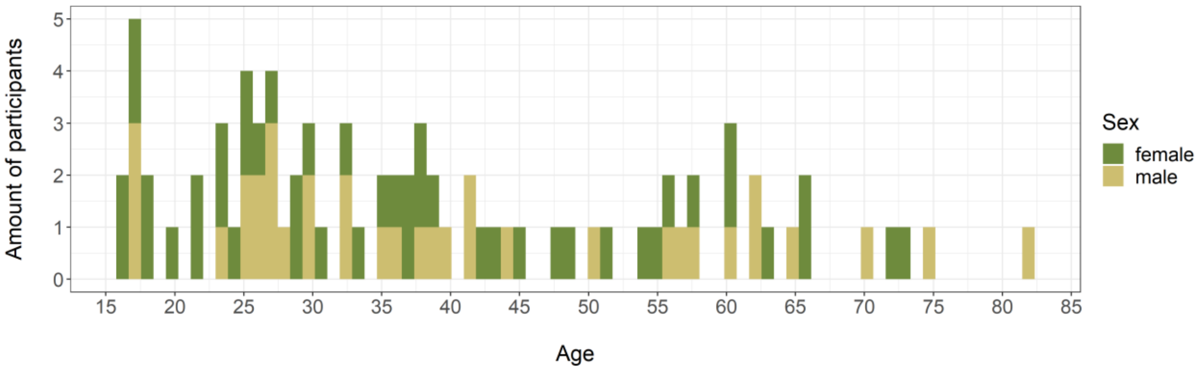


Table 1

Summary of demographic participant data

	n	%
Sex		
female	46	57
male	35	43
Current residence		
Italy	79	98
other country	2	2
Additional first languages		
no	76	94
yes	3	4
not specified	2	2
Dialect		
no	51	63
yes	29	36
not specified	1	1
Knowledge of second languages		
yes	66	82
no	14	17
not specified	1	1
Musical experience		
yes, instrument	28	34
yes, singing	8	10
yes, both	19	24
no	26	32

3.2. Data collection

3.2.1. Pre-experimental procedure

Prior to data collection, participants were informed about the experimental procedure as well as data protection measures. They were then given the opportunity to ask any questions before signing a consent form. All participants completed a questionnaire inquiring the demographic data presented in Table 1. Subsequently, each participant underwent a hearing assessment using a Maico MA25 audiometer to ensure typical hearing thresholds for their age. During this assessment, pure tones of 0.25, 0.5, 1, 2 and 4 kHz were tested at increasing amplitudes. Participants were instructed to indicate when they were able to hear the tone, thereby revealing their hearing thresholds for the different frequencies. Typical hearing related to individuals' age was based on Schow (1991). This corresponded to the following criteria: up to 50 years of age, hearing threshold cut-offs for all frequencies were 25 dB HL. Above 50 years of age, hearing threshold cut-offs at 250 and 500 Hz were 25 dB HL, while hearing thresholds cut-offs for frequencies of 1 Hz and above were 40 dB HL. The hearing screening took about 2 minutes per participant.

3.2.2. Experimental setup

All participants completed an experimental procedure consisting of three parts: (1) reading a text passage ("L'arcobaleno") aloud, (2) an auditory-motor adaptation experiment using altered auditory feedback, and (3) reading the same text passage ("L'arcobaleno") aloud again. Overall, the experimental procedure took approximately 15 minutes. This master thesis will focus on a subset of the experimental procedure: the auditory-motor adaptation experiment using altered auditory feedback.

The auditory-motor adaptation paradigm: Auditory feedback perturbation. For the experimental feedback perturbation task, participants were equipped with an over-the-ear microphone (i.e., Shure MX153 T), which was placed 7 cm away from the mouth and adjusted to a 45-degree angle. They were also fitted with headphones (Sennheiser HD 280 Pro). The participants were instructed to produce the word they saw on the screen in front of them when prompted. They were asked to produce one of three disyllabic Italian words: /ˈbeːve/, /ˈveːde/, or /ˈdeːve/ (meaning 'drink', 'see', 'must'). These words were chosen so that: (1) they adhere to the syllable structure CVCV since etymologically native Italian words tend to prefer open over closed syllables at word final positions (Grice et al., 2015; Miatto, 2020), (2) the vowel remained the same after the first and second consonant, and (3) when the first formant (F_1) in /e/ was shifted, all words remained 'real' words. That is, the words became similar to 'bava', 'vada', and 'dava', which translate to 'drool', 'go', and 'gave'. Participants were instructed to prolong the first vowel /eː/ compared to their regular speaking habits in order to achieve 1 – 2 s of vowel production for perturbations of the F_1. The disyllabic target words were prompted by presenting them in a black font on the white background of a screen placed in front of them. For practice, the participants produced nine words without any perturbation of the feedback perceived through the headphones. That is, each target word was prompted three times in random order per phase of the experiment. Overall, each participant produced 108 target words throughout the experimental task. As participants were speaking, the produced F_1 of the stressed vowel /eː/ was manipulated in a stepwise manner (Houde & Jordan, 1998; Kearney et al., 2020). This manipulation was implemented by the software Audapter 2.1.012 (Cai et al., 2008), which was run in MATLAB R2019a. Participants' productions with or without altered F_1 was sent back to them via headphones and a Focusrite Scarlett Solo (2nd Gen) soundcard. More precisely, F_1 perturbation in participants' auditory feedback was conducted in four phases: baseline, ramp, hold, and after-effect.

1) **Baseline phase.** The initial phase spanned over 24 trials. Participants received their typical auditory feedback via headphones, while they were speaking. This phase provided the individual baselines for the first formant of /eː/.
2) **Ramp phase.** Throughout the following 30 trials, the F_1 in participants' auditory feedback was gradually increased. In each subsequent trial, the F_1 was 1.7 % higher than in the previous trial.
3) **Hold phase.** For the subsequent 30 trials, the perturbation of F_1 in the vowels of the stimuli was held constant at a 50 % increase relative to the average F_1 of the individuals' baseline trials. The maximum of 50 % in F_1 shifting was determined through piloting of the presented experiment.
4) **After-effect phase.** For the last 24 trials, there were no perturbations. Typical auditory feedback was delivered via headphones the same way as during the baseline phase. That is, to study the immediate after-effect of the applied feedback perturbation, the final phase did not involve F_1 perturbations.

3.3. Data analysis

3.3.1. Pre-processing

Pre-processing of the data was based on previous research using this experimental paradigm (Abur et al., 2021; Lester-Smith et al., 2020; Van Brenk & Terband, 2020). F_1 vectors of participants' productions of the first stressed vowel /eː/ in each word were estimated by Audapter during data collection by means of linear predictive coding (LPC) and a sampling rate of 16,000. Using MATLAB 2013b and a custom-written script, those F_1 vectors were subsequently extracted. Each vector was plotted across time on a wideband spectrogram with a window length of 0.005 s and 0 to 5,000 Hz in order to be manually inspected for accuracy. In the case of inappropriate automatic selection of the first /eː/, the correct F_1 window was selected manually.

The custom MATLAB script then calculated the mean F_1 value of each stressed vowel production from a segment of 40 to 120 ms of the selected time window for the vowel to estimate feedforward control. At the onset of that window, the F_1 has previously been found to stabilize, while after that window, there are thought to be greater contributions from the feedback control system (Abur et al., 2021; Lester-Smith et al., 2020). In the course of this pre-processing stage, trials were excluded if no vocalization occurred in the recording due to high response latencies or background noise disrupted the vocalization (n = 38).

Lastly, mean F_1 values across all trials were normalized as percent deviation from the mean F_1 of the baseline phase as shown in Equation 1. The baseline phase comprises 24 trials without perturbed feedback, thereby providing an indication of individual participants' typical range in F_1 during productions of the target vowel. The outcome value of this equation is herein referred to as normalized F_1. Normalized F_1 values of each trial's target vowel were used as the dependent variable in the statistical analysis.

Equation 1

Normalization of vowels: Percentage change from the baseline

$$normalized\ F_1[n] = \frac{(mean\ F_1[n] - \mathrm{mean\ F_1}[1{,}24]) * 100}{mean\ F_1[1{,}24]}$$

Note. n indicates the number of a given trial in an experimental session (i.e., 1 – 108).

Additionally, the custom MATLAB script extracted all recordings altered in F_1 that were perceived by the participant during the experiment (i.e., the headphone signal). During data pre-processing, these recordings were manually checked for acoustic quality using Praat (Boersma & Weenink, 2023). No trials had to be excluded due to poor quality in the course of this inspection.

3.3.2. Statistical analysis

The data were statistically analysed by fitting generalized additive mixed models (GAMM) by means of R version 4.2.2 in RStudio (R Core R Core Team, 2022). The choice of this method was based on the non-linear nature of the present data, and speech data in more general terms. GAM is a regression method that allows the researcher to identify patterns over time, while accounting for autocorrelating residuals in the data (Wieling, 2018). Model fitting was based on discretization of covariate data (Li & Wood, 2020; Wood et al., 2017). GAMMs were fitted and visualized using the R packages *itsadug* (Van Rij et al., 2022) and *mgcv* (Wood, 2011). For modelling specifically, the function *bam* from the latter package was used.

To assess adaptive responses across trials, normalized F_1 values per trial were used as the dependent variable. The final model per dependent variable was selected in a stepwise procedure starting from a basic model of F_1 trajectories over time (i.e., trial number 1 to 108). This selection is described in more detail in the results section. For increased comparability to previous work, averages of individual participants' normalized F_1 values in the hold phase were calculated additionally and plotted using *ggplot2* (Wickham, 2016). These values are considered the response magnitude as they reflect the maximum percentage change from the individual baseline mean F_1.

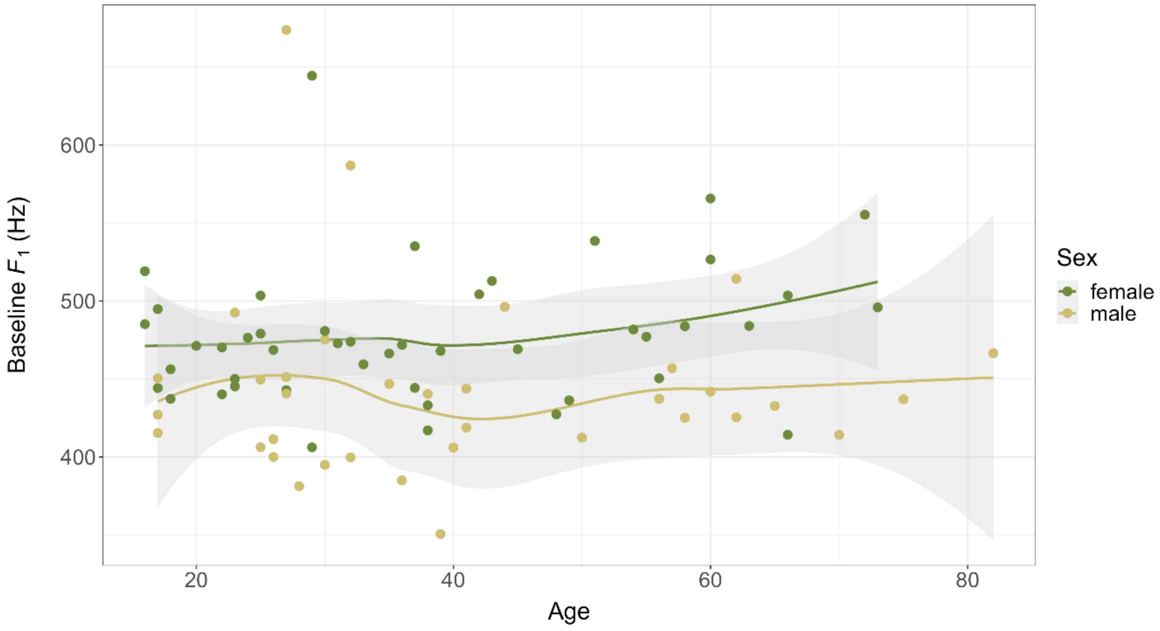
4. Results

This study aimed to contribute to our comprehension of the impact of typical aging on auditory-motor adaptation for articulation. To address this objective, a speech production experiment was conducted wherein the first formant (F_1) in participants' auditory feedback was systematically increased. The experiment was divided into four distinct phases of manipulation: baseline, ramp, hold, and after-effect. The analysis started with general inspection and exploration of the data, followed by hypothesis testing statistics.

In a first step, the data was broadly inspected and described. Female participants showed an average F_1 value of 478 Hertz (Hz) during the baseline condition with a standard deviation (SD) of 44 Hz. Male participants demonstrated a baseline mean F_1 value of 443 Hz with an SD of 58 Hz. Figure 2 shows mean F_1 values per participant based on the productions in the baseline phase of the experiment. The individual mean values plotted in this graph are listed in Appendix 1.

Figure 2

Individual participants' mean baseline F_1



Note. Mean F_1 per participant in the baseline phase (24 trials), values of each trial measured in a time window of 40 to 120 ms within the first vowel in consonant-vowel-consonant-vowel (CVCV) sequences; coloured line represent the means per sex, shaded areas represent 95 % confidence intervals.

4.1. Exploratory analysis

In order to examine the process of articulatory adaptation triggered by the experimental manipulation statistically, first, a basic Generalized Additive Mixed Model (GAMM) was fit to the data. This model is described in Equation 2. The presented formula delineates three smoothing terms, which together model the trajectories of normalized F_1 values over time (i.e., trial 1 to 108) while accounting for interindividual variability between participants and stimuli (i.e., 'beve', 'vede', 'deve'). More specifically, Equation 2 comprises an intercept β_0, smooth functions $f()$ which describe different trajectories over time in individual participants, a random effect $b()$, and an error term ε.

Equation 2

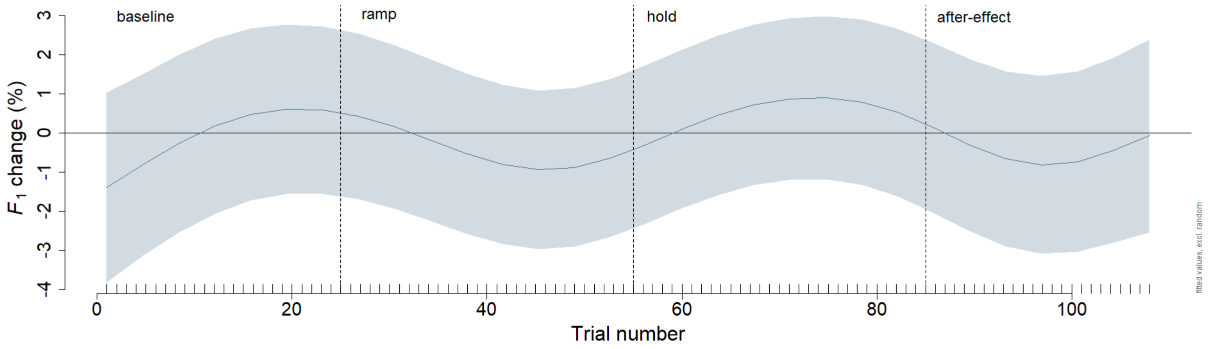
Basic GAMM

$$F_1\, normalized \;=\; \beta_0 \;+\; f_1(trial) \;+\; f_2(trial, participant) \;+\; b(stimulus) \;+\; \varepsilon$$

This basic model was expected to show an overall adaptive response opposing the direction of the F_1 perturbation (Houde & Jordan, 1998; Lester-Smith et al., 2020; Mitsuya et al., 2015; Villacorta et al., 2007). For the present study, the F_1 in the auditory feedback was increased relative to participants' productions in the baseline phase. Accordingly, a gradual decrease in the F_1 of vowel productions was expected. However, this prediction was not confirmed by the initial inspection of the collected data. Figure 3 visualizes the outcome of the basic GAMM described in Equation 2, while the results are additionally presented in Table 2. This basic model suggests a significant non-linear effect of trial ($p < .001$). That is, participants seemed to deviate from their individual mean F_1 of the baseline phase throughout the course of the experiment.

Figure 3

Basic GAMM across all participants



Note. Percentage change of F_1 over the course of the experiment relative to participants' individual baseline means, 95 % confidence intervals are represented by the shaded areas.

Table 2

Results of the basic GAMM

Parametric coefficients	β	SE	*t*	*p*
Intercept	-0.10	0.97	-0.11	0.92
Smooth terms	edf	Ref.df	*F*	*p*
s(trial)	7.28	7.84	4.92	< 0.001
s(trial, participant, bs='fs', m=1)	568.99	728.00	18.921	< 0.001
s(stimulus, bs = 're')	1.99	2.00	290.50	< 0.001

Note. SE = standard error, edf = effective degrees of freedom, Ref.df = reference degrees of freedom, re = random effect, fs = factor smooth; results of the model provided in Equation 3 after applying model criticism.

Due to the unexpected results of data inspection, the data was explored in more detail prior to testing the initial hypothesis. It was found that this pattern had emerged as a result of an almost equal amount of participants opposing (*n* = 41) and following (*n* = 40) the perceived F_1 increase in their productions. Thus, the direction of response in individuals' productions was included as a factorial predictor in the model. The resulting GAMM is shown in Equation 3.

Equation 3

GAMM accounting for response direction

$$F_1\ normalized = \beta_0 + f_1(trial, response\ direction) + \beta_1 * response\ direction + f_2(trial, participant) + b(stimulus) + \varepsilon$$

Due to a skewed distribution of the residuals in the above model, model criticism was applied deleting data points that are more than 2.5 standard deviations away from the mean (i.e., 198 data points). There was no autocorrelation observed in the residuals of this model. Thus, the trimmed data set met the statistical assumptions. The results of the trimmed model are shown in Table 3.

Table 3

Results of the GAMM including response direction

Parametric coefficients	β	SE	*t*	*p*
Intercept	3.03	0.97	3.14	0.002
Response direction: Opposing	-6.39	0.70	-9.15	< 0.001
Smooth terms	edf	Ref.df	*F*	*p*
s(trial): Following	7.85	8.29	12.23	< 0.001
s(trial): Opposing	5.85	6.45	5.39	< 0.001
s(trial, participant, bs='fs', m=1)	571.85	727.00	14.53	< 0.001
s(stimulus, bs = 're')	2.00	2.00	402.76	< 0.001

Note. SE = standard error, edf = effective degrees of freedom, Ref.df = reference degrees of freedom, re = random effect, fs = factor smooth; results of the model provided in Equation 3 after applying model criticism.

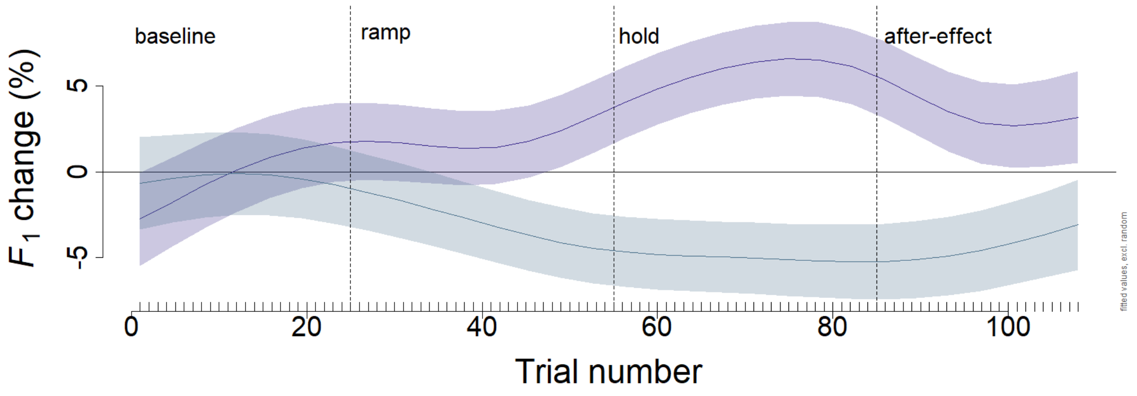
The model as described in Table 3 is visualized per response direction in Figure 4. The statistical analysis revealed a significant difference (β = -6.39, *p* < .001) between two response directions: participants increasing (following) and participants decreasing (opposing) their F_1 productions. The deviation from the baseline mean F_1 in a non-linear manner was significant in both, participants opposing (*p* < .001) as well as participants following the perturbation (*p* < .001).

For participants following the F_1 perturbation (visualized in purple in Figure 4), the observed trajectory of F_1 change overall suggests an initial instability of F_1 productions in the baseline phase. That is, F_1 values of baseline trials appear to deviate more strongly from the baseline mean than in the group of participants opposing the perturbation. A clear following response to the perceived F_1 increase starts during the ramp phase and continues in the hold phase. Participants returned to the baseline during the after-effect phase. This plot pertains to the first smooth term in Table 3, i.e., *s(trial): Following*.

On the contrary, participants who opposed the F_1 perturbation (visualized in grey in Figure 4) demonstrated comparatively stable F_1 productions during the baseline phase. An opposing pattern could be observed during the ramp phase, and tended to stabilize during the hold phase. The grey line in Figure 4 illustrates the second smooth term in Table 3, i.e., *s(trial): Opposing*. The temporal incongruency between the initiation of opposing responses and the initiation of following responses (see ramp and hold phases in Figure 4) seemed to have caused the wave pattern observed in Figure 3 above.

Figure 4

GAMM smoothed terms for response directions



Note. Visualization of smooth terms by response direction presented in Equation 3 and Table 3, grey line = *s(trial): opposing*, purple line = *s(trial): following*; x-axis shows percentage change of F_1 over the course of the experiment relative to participants' individual baseline means, 95 % confidence intervals are represented by the shaded areas.

Finally, a Wilcoxon Rank Sum test was run to specifically compare baseline F_1 values of participants following and participants opposing the perturbation. This control step was added in order to examine whether different baseline productions due to allophonic variability in the /e/ phoneme between speakers may be involved in the difference in response directions. Raw F_1 values from the baseline phase (i.e., not normalized as described in the Methods section) averaged per participant (shown in Figure 2 and Appendix 1) were used as the dependent variable. However, there was no significant difference in the baseline F_1 between the group of opposers and the group of followers ($W = 832$, $p = 0.914$).

4.2. Hypothesis testing

The initial hypothesis of this thesis states that aging has a sex-dependent negative impact on auditory-motor adaptation for articulation. This hypothesis was tested through building a GAMM based on the preceding exploratory analysis. Age was included as a continuous predictor and sex as a factorial predictor. Neither variable had a significant effect as shown in Table 4. The hypothesis testing model is described in Equation 4. It includes tensor product interactions (i.e., *ti()*) between the trial number (i.e., the experimental time course) and the variable age, which can vary by the levels of response direction or sex. Similar to the previous GAMM, model criticism was applied due to non-normally distributed residuals and highly influential outliers were excluded.

Equation 4

Hypothesis testing GAMM

$$F_1\ normalized = \beta_0 + f_1(trial, response\ direction) + f_2(age, \mathrm{r}esponse\ direction) + ti_1(trial, age, response\ direction) + \beta_1 * response\ direction + f_3(age, sex) + ti_2(trial, age, sex) + \beta_2 * sex + f_4(trial, \text{participant}) + b(stimulus) + \varepsilon$$

The model in Equation 3, explained 60 % of the deviance in the data. The results of the hypothesis testing model as listed in Table 4 show a significant effect of response direction (i.e., opposing vs. following) alongside significant random effects structures. The main effect of response direction showed that participants opposing the F_1 shift differed significantly from participants following the feedback shift (β = -5.32, $p < .001$) with respect to their F_1 productions. Both groups of participants, those following ($F = 11.65$, $p < .001$) and those opposing ($F = 4.76$, $p < .001$) the perturbation, significantly adapted their F_1 productions in the course of the experiment. The full model was plotted and presented in Figure 5. It suggests an overall opposing response to the experimental F_1 perturbation.

Table 4

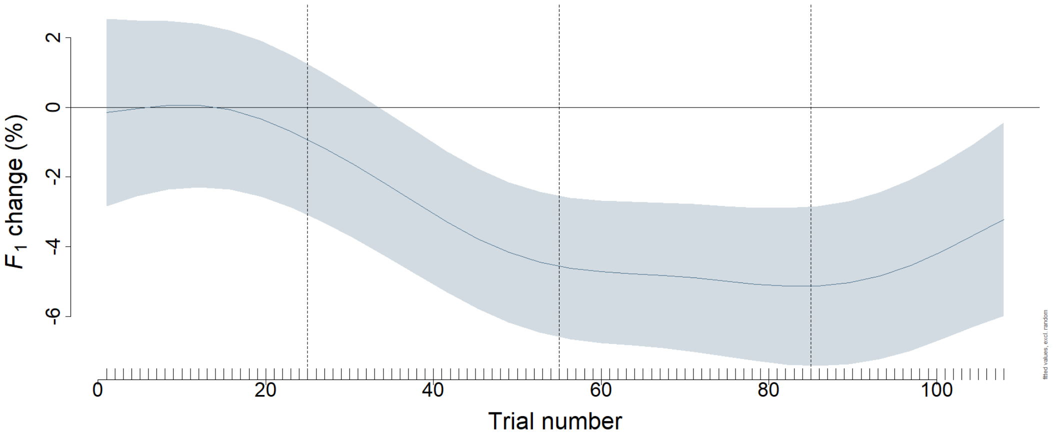
Results of the hypothesis testing GAM

Parametric coefficients	β	SE	*t*	*p*
Intercept	2.41	0.95	2.54	0.011
Response direction: Opposing	-5.32	0.52	-10.21	< 0.001
Sex: Male	-0.07	0.52	ß0.14	0.890
Smooth terms	**edf**	**Ref.df**	***F***	***P***
s(trial): Following	7.86	8.35	11.65	< 0.001
s(trial): Opposing	5.47	6.12	4.76	< 0.001
s(age): Following	1.69	1.71	0.90	0.504
s(age): Opposing	1.00	1.00	0.52	0.471
ti(trial, age): Following	1.00	1.00	0.04	0.840
ti(trial, age): Opposing	2.54	2.71	0.75	0.644
s(age): Female	0.00	0.00	0.01	0.999
s(age): Male	1.00	1.00	0.02	0.897
ti(trial, age): Female	0.00	0.00	0.10	0.991
ti(trial, age): Male	1.00	1.00	0.07	0.799
s(trial, participant, bs='fs', m=1)	516.40	714.00	7.35	< 0.001
s(stimulus, bs = 're')	1.99	3.00	192.33	< 0.001

Note. SE = standard error, edf = effective degrees of freedom, Ref.df = reference degrees of freedom, re = random effect, fs = factor smooth; results of the model described in Equation 4

Figure 5

Visualization of the hypothesis testing model

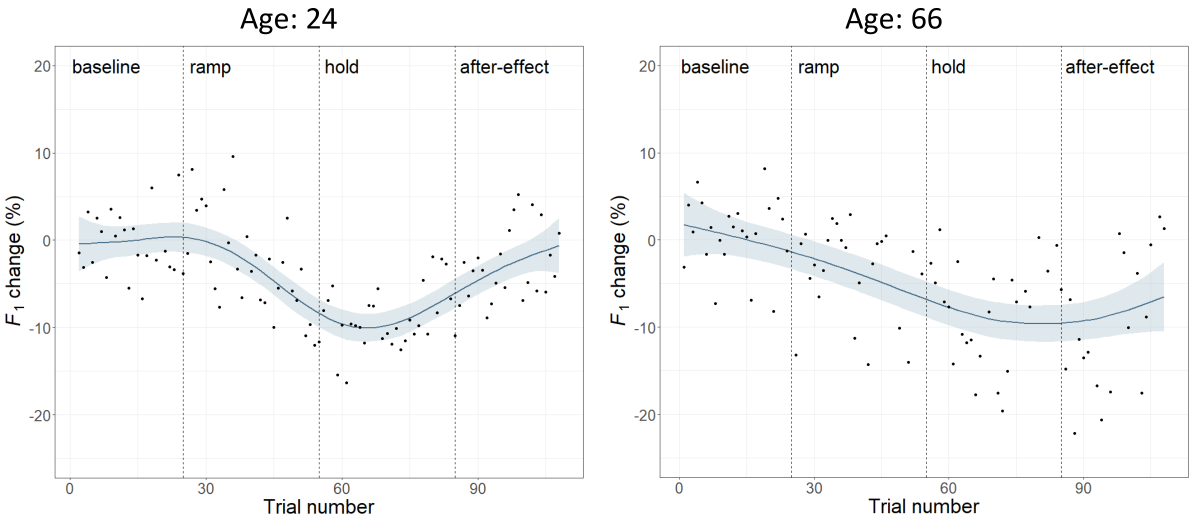


Note. Visualization of the model provided in Equation 4; x-axis shows percentage change of F_1 over the course of the experiment relative to participants' individual baseline means, 95 % confidence intervals are represented by shaded areas.

An inspection of individual participants' F_1 trajectories over the time course of the experiment confirmed that both opposing and following responses are seen across age. Figure 6 and Figure 7 exemplify four participants: Two young and two older participants, one of which opposes and one of which follows the perturbation, respectively. While these four individual plots approximate the overall patterns of following and opposing participants, they show between-speaker variability in the exact trajectories of F_1 adaptation. Especially the examples for following responses shown in Figure 7 differ between each other. The younger participant (see Figure 7, right plot) started by opposing the F_1 perturbation in their productions yet reversed the response direction during the ramp phase leading to a temporally shifted positive peak (i.e., a following response). The older participant (see Figure 7, left plot), on the other hand, started adapting their vowel productions in the hold phase, thus with a higher latency than the younger participant.

Figure 6

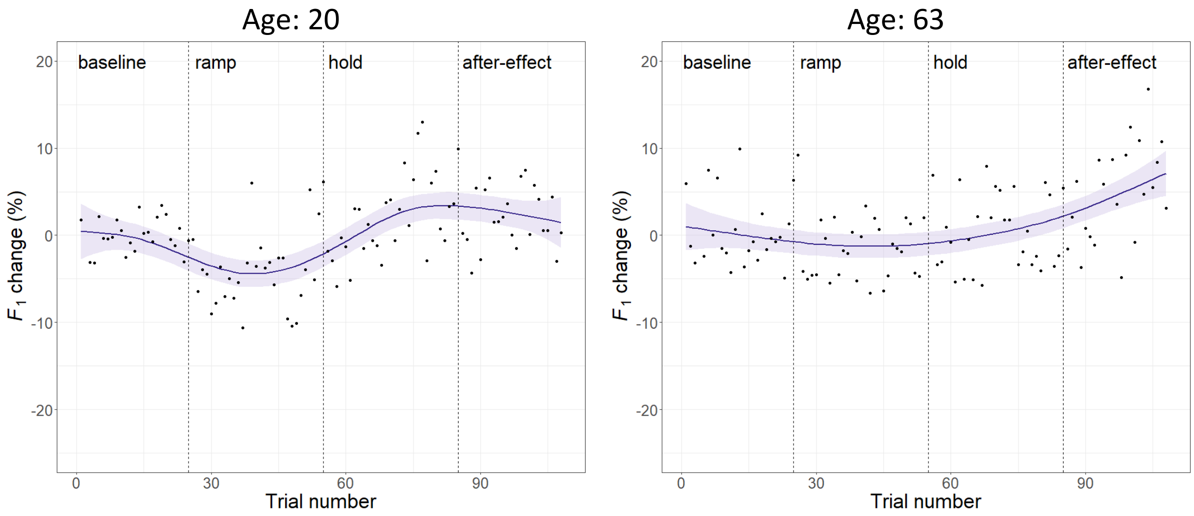
Examples for individual participants opposing the perturbation



Note. Percentage change of F_1 over the course of the experiment relative to participants' individual baseline means, 95 % confidence intervals are represented by shaded areas; black dots correspond to individual normalized F_1 values per trial.

Figure 7

Examples for individual participants following the perturbation

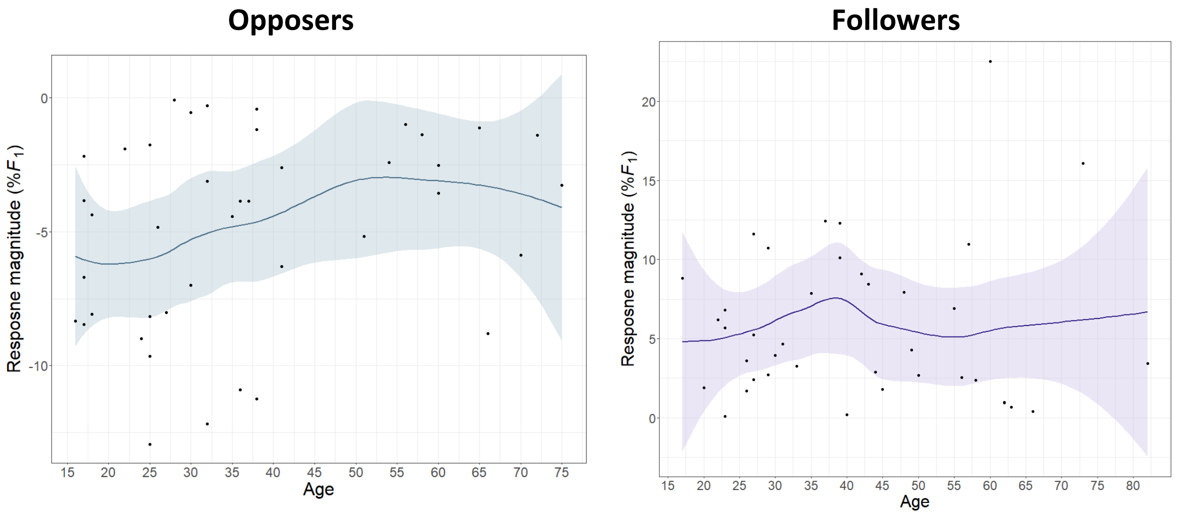


Note. Percentage change of F_1 over the course of the experiment relative to participants' individual baseline means, 95 % confidence intervals are represented by shaded areas; black dots correspond to individual normalized F_1 values per trial.

Lastly, individual response magnitudes were examined for increased comparability to previous research. Response magnitude refers to the mean of normalized F_1, thus deviation from the baseline mean, in the hold phase. Figure 8 presents the relation of age and response magnitude for participants opposing the perturbation (see Figure 8, left plot) and participants following the perturbation (see Figure 8, right plot). The corresponding response magnitudes per participant are listed in Appendix 2. Among the participants who opposed the perturbation, response magnitudes ranged from -13 % to -0.09 %. Among those following the perturbation, response magnitudes ranged from +0.10 % to +23 %. In order to assess correlations between age and response magnitudes statistically, Spearman's rho was estimated. However, there was no significant correlation between age and response magnitude, neither in the group of participants opposing the perturbation (ρ = 0.297, p = 0.059), nor in the group of participants following the perturbation (ρ = -0.80, p = 0.630).

Figure 8

Response magnitudes per participant plotted against age


Opposers

Followers

Note. Response magnitude = mean percentage deviation of F_1 in the hold phase compared to the baseline phase per participant; black dots correspond to individual participants, coloured vertical lines correspond to the correlation of response magnitude and age, shaded areas correspond to 95 % confidence intervals.

5. Discussion

This thesis revolved around the central question of how typical aging impacts auditory-motor control, specifically with respect to articulatory adaptation. The hypothesis posited that auditory-motor adaptation would decline with age, exhibiting distinct patterns between sexes. To test this hypothesis, a perturbation was introduced by shifting the first formant (F_1) in participants' auditory feedback, which corresponds to the height of the tongue in the oral cavity. Consequently, motor adaptation of tongue height was expected as a response to this feedback perturbation. The measure of motor adaptation was indirectly assessed by calculating the percentage deviation of F_1 from each participant's individual baseline mean F_1 in their subsequent productions. It was anticipated that the magnitude of this deviation, in response to the experimental feedback perturbation (i.e., F_1 increase), would increase with age. In this thesis, F_1 change was examined as a gradual process across the experimental trials. Additionally, the participants' average F_1 deviation during the phase of maximum perturbation (i.e., the response magnitude in the hold phase) was computed to facilitate comparison with previous research, which predominantly reports this specific value as the adaptation response magnitude. However, the findings of this study did not support the initial hypothesis, as neither a significant effect of age in relation to the gradual perturbation nor a significant interaction effect of age and sex were observed in the presented data.

Unexpectedly, during an exploratory analysis, the current results unveiled substantial interindividual variability in participants' responses to auditory feedback perturbations. Specifically, a significant effect of response direction was observed: 40 participants followed the perceived increase of F_1, while 41 participants exhibited an opposing response. These two groups of participants did not show any significant differences in their baseline productions of F_1. Consequently, this suggests that the response direction was not significantly influenced by the allophonic variants of /e/ utilized by individual participants. This split in response directions did also not appear to be linked to aging.

5.1. Directions of adaptive responses: Following vs. opposing the perturbation

In response to the experimental F_1 increase in the auditory feedback, approximately half of the participants decreased the F_1 in their subsequent productions. This opposing response is in line with the initial expectation for the overall production patterns over time and thus, the response magnitude in the hold phase. According to the DIVA model (Tourville & Guenther, 2011), speech motor programmes, which are associated with specific acoustic outputs (i.e., phonemes of a given language), are stored in the speech sound map. While these representations are developed during childhood, they remain adaptable in adulthood (Cai et al., 2023; Houde & Jordan, 1998; Scheerer et al., 2016; Villacorta et al., 2007). Therefore, if the auditory feedback consistently deviates from a target sound, the corresponding speech motor programme undergoes adaptation. Consequently, in the aforementioned experiment, it was anticipated that unexpectedly high F_1 frequencies in the auditory feedback would be countered by raising the tongue higher to generate a lower F_1. Hence, it appears that approximately half of the participants attempted to revert to their baseline production of /e/ by lowering the F_1 in their speech output while receiving deviating auditory feedback. This mechanism of auditory-motor adaptation for vowel articulation has been empirically supported by many previous studies employing similar paradigms (Houde & Jordan, 1998; Mitsuya et al., 2015; Villacorta et al., 2007).

Furthermore, roughly half of the data presented in this thesis does not coincide with this finding. Contrary to expectations, many participants exhibited a response in line with the perceived increase of F_1, resulting in an increase in their F_1 productions following the feedback manipulation. This unexpected response challenges the anticipated mechanism described earlier. Although the participants who followed the perturbation significantly adapted their speech motor programmes in response to the manipulation, they did not appear to attempt to revert to their baseline productions. However, there seems to be a temporal difference between the two groups concerning the onset of F_1 changes. Participants who opposed the perturbation exhibited changes in their F_1 productions that aligned with the perturbation, starting from the beginning of the ramp phase onwards. Conversely, participants who followed the perturbation seemed to initiate the response to the perturbation later during the ramp phase. The current analysis did not suggest a significant role of aging for the direction of adaptive responses. Similarly, aging did not correlate significantly with the magnitude of following responses. The reasons behind the high ratio of following responses evident in the outcome are yet unclear but this discussion provides potential explanations beyond age. These explanations primarily focus on allophonic variability of the target vowel /e/ in the Italian language (see section 5.1.1. *Producing the Italian /e/: Allophonic variability and phoneme boundaries*) and the structure of the experimental stimuli used in this study (see section 5.1.2. *The structure of the experimental stimuli*).

Similar to the findings presented above, Kothare et al. (2020) found following alongside opposing responses to formant shifting in young adults (mean years of age = 28.83, SD = 11.82). The authors shifted F_1 and the second formant (F_2) in participants' auditory feedback from /ɛ/ in six different directions within the vowel space. They employed two different experimental approaches: using equal shifting magnitudes (50 Hz) across all participants, and individually adapting the magnitude of formant shifting based on individual distances within the vowel space to make sure that each participant heard categorically different vowels. Shifting in the direction of /æ/ and /ɑ/, which is similar to the present study, overall led to following responses when the magnitude of shifting (50 Hz) was the same for all participants. Conversely, when the magnitude of shifting was individually adapted to each participant's distances between phonemes in the vowel space, the authors observed opposing responses overall. In the present study, the magnitude of shifting was calculated on an individual percentage basis (up to +50 %). This approach allowed for individualization of the perturbation for each participant and thus accounted for allophonic variability of the vowel /e/ (e.g., due to regional variability).

5.1.1. Producing the Italian /e/: Allophonic variability and phoneme boundaries

As suggested by the study conducted by Kothare et al. (2020) , individual variability in the distribution of vowels within the vowel space, encompassing variability within phonemes and distances between phonemes, may contribute to individuals' adaptation to auditory feedback perturbation. This variability in the vowel space is evident in both speech production and perception. Notably, more stable (i.e., less variability within phoneme categories) and more acoustically disparate phoneme production correlates with more precise perceptual discrimination of speech sounds (Franken et al., 2017; Perkell et al., 2004). In turn, Villacorta et al. (2007) proposed that individuals with more precise perceptual discrimination abilities have smaller phonemic target regions, as inferred from a theoretical interpretation of their study findings. In their investigation, they examined young adult native speakers of English (median years of age = 21) with regards to auditory discrimination of stimuli that varied in F_1, as well as auditory-motor adaptation to both upward and downward shifts in F_1. Their results demonstrated that individuals with more precise discriminative abilities exhibited magnified behavioural responses to perturbations. For the present study, however, acuity of vowel perception was not assessed.

Allophonic variability and phoneme boundaries within individuals' vowel spaces may additionally be influenced by regional accents. While all participants in the present study were native speakers of Italian, the Italian language exhibits considerable variability of accents. Especially the mid vowels /e, ɛ, o, ɔ/ strongly differ regionally (Crocco, 2017). Thus, participants of the current study may differ regarding the variability allowed within their individual phoneme target regions on the speech sound map, leading to differences in the distance between vowels of their vowel space. Both within- and between-vowel variability might have influenced the response patterns. In general, auditory feedback manipulation approaching a different phoneme more closely tends to trigger stronger responses (Niziolek & Guenther, 2013). Therefore, individual F_1 values of baseline /e/ productions were reviewed in more detail. In the baseline trials of the current data, which did hence not involve feedback perturbation, allophonic variability of the F_1 values for /e/ largely ranged between 400 and 500 Hz. Similarly, Esposito (2002) reported the mean F_1 for Italian /e/ to be 439 Hz after labial and 433 Hz after alveolar consonants (herein: 'beve', 'vede', 'deve'). However, standard Italian differentiates two mid-front phonemes: /e/ and /ɛ/ (Rogers & d'Arcangeli, 2004), which are not necessarily distinguished in all Italian dialects (Crocco, 2017). The latter is pronounced with an average F_1 of 526 Hz after labials and 464 Hz after alveolars (Esposito, 2002). The auditory perturbation might thus have closely approached or even crossed categorical boundaries for some participants of the present study but not for others. Corresponding differences in the auditory target map might have contributed to different response patterns. While interindividual phonetic differences in the baseline F_1 values of /e/ productions did not prove to statistically predict the adaptive responses in the current study, the experiment reported here did not include any measurements of vowel perception or discrimination. However, such additional information could deepen the insights into the relation between the variability of speech targets and adaptation. Additionally, closed vowels such as /e/ have been demonstrated to induce smaller adaptive responses than more open vowels such as /ɛ/ (Mitsuya et al., 2015). The openness of a vowel is determined by the F_1. Hence, this previous finding further substantiates the notion that individual variants of /e/ among the current participants could be potential factors influencing their adaptive responses.

While within-phoneme variability and surrounding vowels on the speech sound map seem to influence adaptation of a given target vowel, the existing literature is inconclusive regarding the exact effect of those neighbours on the magnitude of auditory-motor adaptation. In the study conducted by Villacorta et al. (2007), response magnitudes differed for upward and downward manipulation (+/-30 %), which, according to the authors, might be attributed to the varying distances from categorical phoneme boundaries. Participants whose feedback was decreased opposed the perturbation to a lesser degree than participants whose feedback was increased. The authors don't report on the possibility of participants following the perturbation. However, a higher ratio of individuals following the downward than following the upward perturbation might have reduced the overall response magnitude for downward shifting in comparison to upward shifting. On the other hand, Franken et al. (2019) reported that speakers tended to oppose downward shifts in F_1 more strongly than upward shifts. Notably, Villacorta et al. (2007) examined native speakers of English, whereas Franken et al. (2019) investigated native speakers of Dutch, indicating differences in the vowel spaces of their respective participants. Likewise, the vowel spaces of the Italian native speakers involved in the current study most probably differed from those in the previous studies, which likely influenced the response magnitudes observed.

Thus, differences in vowel spaces resulting from participants' specific language backgrounds appear to influence adaptive responses to formant shifting in auditory feedback. Accordingly, allophonic variability due to substantial regional diversity in Italian speakers (Crocco, 2017) was thought to potentially have led to different response patterns in a similar vein. Kothare et al. (2020), who similarly reported following besides opposing responses, included participants who are multilingual or monolingual fluent, yet not exclusively first language (L1) speakers of English. This variety of language backgrounds in their sample of participants might have entailed an increase in baseline variability of vowel formants (Cai et al., 2023) in the research by Kothare et al. (2020).

Another potential factor contributing to the unexpected ratio of opposing and following responses could be of a methodological nature. The maximum feedback shift in the current experiment was a 50 % increase relative to the participants' individual mean F_1 in the baseline phase. This was higher than in previous studies that involved relative adaptations as compared to the individual baseline, which commonly used 30 % (Abur et al., 2021; Lester-Smith et al., 2020; Villacorta et al., 2007). Feedback shifting that is too strong might not be interpreted as self-induced and thus not be opposed but rather understood as a different phonemic target (Daliri & Dittman, 2019). However, the latter studies, among many other previous articulatory adaptation studies, were conducted with English-speaking participants, whereas the present study was conducted with Italian-speaking participants. There is no known prior work on auditory-motor adaptation in speakers of Italian. The appropriate perturbation magnitude had therefore been specifically determined via pilot testing prior to data collection. The comparatively high perturbation magnitude required for auditory-motor adaptation in Italian speakers suggested that overall, the participants might perceptually accept high variability for the phoneme /e/. Yet, phoneme distances on the vowel space along with perceptual acuity might differ individually (Franken et al., 2017; Perkell et al., 2004). Thus, the current participants might have differed in the perceived change in vowel categories. Participants following the perturbation might have shown a comparatively delayed start of F_1 changes because of lower auditory acuity. (Delayed) following responses might subsequently have resulted in cases where no perceptual phoneme boundary was clearly crossed.

5.1.2. The structure of the experimental stimuli

The present study differs from most previous work on auditory-motor adaptation in terms of the presented stimuli since there is no prior work on auditory-motor adaptation in Italian available. The experimental stimuli were determined by the phonological characteristics of the target language. As Italian typically does not allow for word-final consonants (Repetti, 2012; Rogers & d'Arcangeli, 2004), the experimental stimuli of this study were existing Italian words consisting of two open syllables, i.e., consonant-vowel-consonant-vowel (CVCV) sequences. Formant perturbation was applied on both vowels. Other studies that employed F_1 shifting, such as those discussed above, studied speech motor control behaviour in response to CVC-structured nonwords in fluent English speakers (Kothare et al., 2020), CVC words in L1 speakers of English (Abur et al., 2021; Mitsuya et al., 2015; Villacorta et al., 2007) or isolated vowels in L1 speakers of Dutch (Franken et al., 2019). Thus, the stimuli of most previous studies, which predominantly reported opposing responses to perturbation, included one target vowel for perturbation. Lametti et al. (2018), on the other hand, investigated auditory-motor adaptation on the level of English sentences in 40 adult L1 speakers (18 to 40 years of age). They applied F_1 increase together with F_2 decrease or vice versa across the vowel spectrum. Their participants not only opposed the perturbation by 30 % in the sentences, but this effect also transferred to single words. The authors did not report on any participants following the formant shifting rather than opposing it. Thus, evidence demonstrates that the feedback perturbation paradigm can induce articulatory adaptation beyond isolated CVC words. This suggests that the structure of the stimuli used in the current study did likely not influence the participants' adaptation behaviour.

5.1.3. Summary of potential confounds

In sum, a variety of factors might influence auditory-motor adaptation resulting in the behaviour observed in the studied data. Auditory-motor adaptation for articulation not only depends on the target vowel itself but may be influenced by the vowel space surrounding a given target. Individually accepted variability within the target phoneme as well as acoustic distance to other vowels in the vowel space might contribute to adaptation in response to perturbation. Individual vowel space configurations, in turn, may be influenced by the regional varieties of a language. Moreover, perceptual acuity and integration of sensory information might influence adaptive response patterns. Structural properties such as length and complexity of the experimental stimuli were not considered to have a substantial effect on adaptation behaviour. However, this possibility cannot be ruled out entirely as a comparison of differently structured stimuli in Italian is out of the scope of the current study.

5.2. Aging

Contrary to the hypotheses of this thesis, the current data did not provide evidence for an impact of age on the magnitude of adaptive behaviour for articulation. That is, age did not have a significant effect on F_1 change throughout the course of the auditory perturbation. Furthermore, age did not correlate significantly with the response magnitude (i.e., individual mean F_1 deviation with maximum perturbation during the hold phase). Employing a paradigm and analysis similar to the present study (i.e., relative F_1 increase), previous research found opposing response magnitudes in the hold phase to be between -18 % and -2 %, with an average of -8.6 % (SD = 6.2 %), in young female participants (18 – 26 years of age; Lester-Smith et al., 2020). In another group of typical adults (48 – 81 years of age), F_1 response magnitudes were -7.2 % on average (SD = 7.6 %; Abur et al., 2021). While this interindividual variability in opposing response magnitudes seems comparable across studies (-13 % to -0.09 % in the results of this thesis), the current analysis did not suggest a correlation of opposing response magnitudes with age. Magnitudes of following responses (+0.10 % to +23 % in the results of this thesis) did likewise not correlate with age. Likewise, directions of perturbation responses (i.e., following vs. opposing the F_1 shift) were not systematically related to age. Previous literature, on the other hand, indicates deterioration of different mechanisms relevant for speech motor control with age (Hu et al., 2023; Li et al., 2018; Tremblay et al., 2018; Tremblay et al., 2017). Specifically, feedback control of pitch as examined through pitch perturbation was found to be impacted by aging (Li et al., 2018; Liu et al., 2011). Besides, behavioural investigations of unperturbed speech production (Hermes et al., 2018; Tremblay et al., 2018) and research on neurophysiological correlates of speech (Tremblay et al., 2017) provided the basis for inferring age-related changes in speech motor control. Still, the complex patterns of aging speech motor control are not fully understood yet (Hu et al., 2023). This thesis added work on aging of auditory-motor adaptation of articulation, a gap in the literature so far. Since the topic of typical aging is novel in the field of speech motor control, no prior results particularly on auditory-motor adaptation based on perturbed feedback are currently available for direct comparison.

However, prior work on typical aging of articulation provides promising complements to the study of auditory-motor adaptation of articulation, adding to a better understanding of aging speech. In particular, aging processes of the speech system do not seem to always be clearly reflected in acoustic measurements. When studying typical, unperturbed articulation in a group of younger (20 – 30 years of age) and a group of older (70 – 80 years of age) adults, Hermes et al. (2018) did not find differences in the acoustic domain. Similar to the present study, they used disyllabic target nonwords of the structure CVCV as target stimuli, but embedded them in German carrier sentences. Measuring articulatory movements directly through electromagnetic articulography, however, Hermes et al. (2018) found reduced peak velocities of the tongue in the older compared to the younger group. In addition, older adults showed significantly longer phases of deceleration after the time point of peak velocity and before the time when the tongue has reached the position for a given target sound. The authors concluded that older adults show slower integration of sensory feedback information for typical speech. This conclusion is based on the notion that the period of deceleration before reaching the target position constitutes the period when sensory feedback is utilized to make any necessary corrections to the movement (Hermes et al., 2018). A prolonged deceleration phase in older compared to younger adults is thus indicative of prolonged sensory integration or speech planning times due to aging. Both sensory integration and speech planning may influence auditory-motor adaptation. Accordingly, kinematic responses to F_1 shifting might reveal a different picture on auditory-motor adaptation than acoustic responses.

Previous work using feedback perturbation paradigms could delineate a systematic aging process of feedback control for voice pitch (i.e., f_o; Li et al., 2018; Liu et al., 2011; Scheerer et al., 2016). More precisely, it was found that magnitudes of opposing responses to sudden f_o perturbations (i.e., reflexive responses) are more variable yet overall smaller in children (3 – 8 years of age) than young adults (19 – 27 years of age; Scheerer et al., 2016). However, in the line of the DIVA model's argumentation, children would be expected to show greater responses to feedback perturbations than (young) adults as their speech system is still developing, and sensory feedback is still required to shape feedforward commands. The latter should theoretically be mainly relied on during adulthood. This hypothesis was not supported by the results of Scheerer et al. (2016). Thus, children might rely on feedforward control for pitch regulation early on. Opposing response magnitudes seem to increase further up to a maximum reached by the fifth decade of life (Liu et al., 2011). From the sixth decade onwards, in turn, perturbation responses tend to decrease (Liu et al., 2011). Still, opposing responses to f_o perturbations were found to be significantly larger in people above 60 years of age compared to younger (19 – 25 years of age) adults (Li et al., 2018). These results could be interpreted as indicating an increase in reliance on feedback control via sensory feedback during adulthood and up to the fifth decade of life, followed by a decrease in reliance on feedback control above the age of 60 years. This behavioural difference is accompanied by differences in terms of timing and magnitude of brain activity (Li et al., 2018). A decrease in response magnitudes to f_o perturbation during later adulthood (60+ years of age) is associated with delayed and reduced early event-related potentials, specifically N1 and P2, compared to younger adults (19 – 25 years of age). A decrease of reliance on sensory feedback in later adulthood might thus be induced by reorganized cognitive resources besides reduced availability of auditory (Slade et al., 2020) and proprioceptive (Ketcham & Stelmach, 2004) information.

On the other hand, adaptive responses to formant shifting, reflecting articulatory control, draw a different picture. Van Brenk and Terband (2020) shifted F_1 along with F_2 in CVC words produced by children (4 – 9 years old) and young adults (18 – 29 years old). In accordance with the DIVA model, they found significantly larger adaptive responses in children compared to young adults, suggesting stronger auditory-motor integration and less stabilized feedforward commands in the former group. The present study was the first to investigate behavioural responses to gradual shifting of F_1 across adulthood, and thereby filled a gap in the existing literature. Based on the presented results, the auditory-motor adaptation mechanism, specifically for vowel articulation, does not seem to change after early adulthood, when feedforward commands have been ingrained.

This study differs from prior work on aging speech motor control in the targeted speech domain (i.e., articulation rather than pitch) and response mechanism (i.e., adaptive rather than reflexive responses). Interpreting the current study in relation to the literature reviewed above consequently raises two questions: Are there distinct aging trajectories of reflexive and adaptative responses to feedback perturbation? Or are auditory-motor control of vocal pitch and articulation inherently different allowing for distinct aging trajectories? With respect to the former question, Franken et al. (2019) compared responses to gradual feedback shifts of F_1 with responses to sudden feedback shifts of F_1. The authors studied long vocalizations of the isolated vowel /e/ within participants. They found larger responses in the adaptive compared to the reflexive condition and concluded that feedback-based corrections and feedforward-based adaptations are two distinct processes. If this distinction is true, there could be different aging processes for reflexive and adaptive auditory-motor control as reviewed above. The difference between the study by Li et al. (2018) and the current study, however, is not restricted to reflexive and adaptive responses. Rather, there could be distinct auditory-motor control processes, with distinct aging trajectories, for voice pitch and articulation. This proposal concurs with Lester-Smith et al. (2020) who compared responses to shifts in F_1 to shifts in f_o and found distinct patterns for articulatory and voicing control. They likewise proposed that the two domains might be regulated by distinct control mechanisms. In this regard, it is possible that different speech motor control subsystems undergo distinct aging processes.

5.2.1. Variability of (adapted) motor programmes

Variability of motor skills has previously been studied across age, as it is considered an indicator for the stability of stored motor commands. While according measures are beyond the scope of the current research, the stability of motor functions might play a role for auditory-motor adaptation of articulation. Based on prior investigations, variability in the fine motor commands required for speech production seems to increase with age throughout adulthood (Hermes et al., 2018; Tremblay et al., 2018; Tremblay et al., 2017). Moreover, general motor learning capacities decline above 60 years of age (Ketcham & Stelmach, 2004). Accordingly, adapted motor commands, especially when adapted temporarily, are likely to be more variable in later adulthood. The analysis performed within this master thesis does not include a systematic examination of variability in baseline speech production or speech motor adaptation. However, speech motor control in older age may be characterized by changes regarding the stability of feedforward motor commands or variability in the learning process, rather than the magnitude of the overarching perturbation response which was explored in the present research.

Aging auditory-motor adaptation might additionally be characterized by increased learning latency. While general motor skills (e.g., balance, handwriting, etc.) can be learned or adapted throughout the lifespan, motor adaptation advances at a slower pace in older (above 60 years of age) than younger (18 – 30 years of age) adults. As a consequence, improvements comparable to younger adults may be seen after a longer period of practice (Ketcham & Stelmach, 2004). In order words, latency of auditory-motor adaptation might be shorter in early compared to later adulthood, as seen in the literature on general motor learning. Increased adaptation latency might be accompanied by increased variability in earlier phases of the learning process.

In typical unperturbed speech productions, aging has been associated with higher variability of motor functioning. Regarding general motor control, the aging process is commonly characterized by slower and less accurate movements (Ketcham & Stelmach, 2004), which is reflected in speech motor control. When performing a series of diadochokinetic tasks with varying sequential and articulatory complexity levels, older adults (mean age = 73.9 years) showed less accurate and slower speech productions in comparison to younger adults (mean age = 27.7 years). The authors inferred articulatory sequencing difficulties (Bilodeau-Mercure & Tremblay, 2016). These behavioural changes were confirmed in later research across age (18 – 83 years of age) using trisyllabic nonwords of varying phonological complexity and syllable frequency in Québec French (Tremblay et al., 2018). The researchers found speed-accuracy trade-offs in the younger but not the older participants. Such findings were interpreted as behavioural indication for declining planning processes of articulatory movements (Tremblay et al., 2018). In favour of that interpretation, Bilodeau-Mercure and Tremblay (2016) did not find age-related muscular degeneration to be related to changes in speech timing. Instead, functional magnetic resonance imaging (fMRI) revealed a relation of age differences in brain activity with acoustically longer and more variable response durations in older (mean age = 68.2 years) compared to younger (mean age = 26.8 years) speakers (Tremblay et al., 2017). Also regarding articulatory movements of the tongue, intraindividual variability is higher in older (70 – 80 years of age) than younger (20 – 30 years of age) adults (Hermes et al., 2018). In sum, aging seems to be related to variability of motor functioning, and potentially speech motor control.

5.2.2. Aging and Sex

In the data presented here, no interaction between aging and sex has been found. Preceding studies on articulation across age with a focus on kinematics did not evaluate differences between sexes (Bilodeau-Mercure & Tremblay, 2016; Hermes et al., 2018; Tremblay et al., 2018; Tremblay et al., 2017). From an acoustic perspective, sex-specific aging of vowel formants seems unsystematic and inconclusive (Torre & Barlow, 2009; Traub-Eichhorn et al., 2018). On the contrary, earlier findings on the aging process of auditory-motor control of the voice pitch reported neurophysiological differences between sexes. More specifically, results from electroencephalography (EEG) in response to sudden pitch (i.e., f_o) feedback perturbation showed smaller N1 and P2 amplitudes for young female (mean age = 21.1 years) than young male (mean age = 22.3 years) participants. This difference was absent in older participants (mean age female: 64.4 years, mean age male: 64.7 years; Li et al., 2018).

Different aging trajectories of f_o have also consistently been found through acoustic measurements (Hu et al., 2023; Torre & Barlow, 2009; Traub-Eichhorn et al., 2018). It thus remains possible that speech production patterns across age differ between sexes only in terms of voice but not in terms of articulation.

5.2.3. Implications for the DIVA model

Aging patterns beyond young adulthood are currently not part of the DIVA model. This study does not provide evidence for a need of including such an aging trajectory for motor control of articulation into the model. Previous research, however, indicates that feedback control of voice changes throughout adulthood (Li et al., 2018; Liu et al., 2011). In line with Lester-Smith et al. (2020), this suggests disparate mechanisms for voice and articulatory control, which are not currently explained by the model. Such distinct mechanisms would allow for an integration of distinct aging patterns per domain.

The current results show that 40 out of 81 participants did not respond to the experimental feedback perturbation as expected based on the DIVA model. That is, instead of opposing the F_1 increase in their subsequent productions, they followed this gradual shift. This effect is not explicable through the DIVA model. Thus, additional factors beyond aging, which were not considered in this experiment and may not be considered by DIVA, must have contributed to this outcome. However, it is unclear what exactly caused the nearly equal distribution of following and opposing response directions in the presented data.

5.3. Limitations and avenues for future research

While the current results do not lend further support to the initially hypothesized age effect, this line of research may be pursued in a multitude of ways. Prospective research can build on the methodology, findings, and limitations discussed herein to contribute to a better understanding of speech motor control across age. Key points that could be addressed in the future include a large and balanced sample of participants, kinematic besides acoustic measurements, interactions with sensory acuity and other mechanisms relevant for speech motor control, and factors triggering opposing or following responses.

Although many participants with similar language backgrounds were recruited for the present study, the inevitably broad range of age might require an even larger sample to draw sound conclusions. In addition, the sample of tested participants was skewed towards younger age, while comparatively few adults above the age of 60 years were tested. Previous research, on the other hand, suggests that a turning point of speech motor control behaviour in the domain of voice pitch can be observed from the sixth decade of life onwards (Liu et al., 2011). Prior work on acoustic and kinematic changes in regard of articulation similarly considers people above 60 years of age, independent of sex, their group of older adults (Hermes et al., 2018; Tremblay et al., 2003; Tremblay et al., 2017). Thus, future research should include more participants in order to equally cover a broad range of age.

The current study investigated auditory-motor control of articulation using a well-established perturbation paradigm and acoustic measurement that allowed for comparison to previous work. While previous work has reported aging effects on articulation which were measured acoustically (Bilodeau-Mercure & Tremblay, 2016; Tremblay et al., 2018), the current study failed to find acoustic changes across age. However, as articulatory measurements can give additional insights and reveal effects unidentified through acoustic measurements (Hermes et al., 2018), investigating auditory-motor adaptation to F_1 perturbation using an articulatory method might disclose formerly unnoticed patterns. Therefore, future research on the aging process of auditory-motor control could benefit from combining both, articulatory and acoustic approaches.

Furthermore, future research could focus on diverse measures of speech motor control beyond the response magnitude. Specifically, this study focused on the magnitude and trajectory of adaptation. However, prior work on general motor control indicates that aging is associated with a difference in motor adaptation latency. That is, older adults (above 60 years of age) require more practice than younger adults (18 – 30 years of age) to reach a similar level of adaptation (Ketcham & Stelmach, 2004). An observed trend for increased variability in adaptation with age led to speculation about adaptation latency in the above discussion. Accordingly, future research could study the latency of auditory-motor adaptation for articulation, a highly specialized and over-learned fine motor skill, in a systematic manner.

The experiment for the current thesis did not include assessments of auditory acuity. As reviewed in the introduction, sensory functions constitute a basis for speech processing and as such, influence the speech system (Panouilleres & Mottonen, 2018; Slade et al., 2020; Tremblay et al., 2003). While typical hearing based on age was verified, participants' acuity of auditory perception was not tested in the presented experiment. However, although sensory acuity to F_1 has been found to be statistically related to F_1 adaptation (Villacorta et al., 2007), this finding failed to be replicated (Lester-Smith et al., 2020). As individuals differ in preferences for auditory or somatosensory feedback (Lametti et al., 2012), a potential relation between sensory acuity to F_1 and F_1 adaptation should still be taken into consideration when perturbing auditory but not somatosensory feedback. In sum, the relation between sensory acuity in the different domains and articulatory adaptation is unclear. Therefore, there is a need for further research studying how auditory and somatosensory functions together impact speech adaptation.

Since this study focused on adaptation of articulation, conclusions about the comprehensive speech motor system are limited to this particular mechanism. Specifically, feedback-based adaptations of feedforward motor programmes were studied. For broader insights into speech motor control, feedback and feedforward control systems could be disentangled in future research. Aging patterns could be studied separately for each control system. Future research could further investigate the possibility of aging in specific steps of the speech motor control process, other than articulatory adaptation. In the articulatory domain, Tremblay et al. (2018) proposed an effect of aging on the planning and execution stages of speech production. Hermes et al. (2018), on the other hand, proposed a deceleration of the comparison between sensory target and state maps (i.e., sensory error computation). Similarly, Li et al. (2018) propose an age-related decline of auditory-motor integration for voice pitch control. Further studies are needed in order to clarify which concrete steps of speech motor control, including auditory-motor adaptation, are comprised by aging.

Finally, the high number of following responses observed in this study is inconsistent with previous studies on auditory-motor adaptation of vowel formants in young speakers (age 18 to 33), which report opposing responses on average (Daliri & Dittman, 2019; Lester-Smith et al., 2020; Van Brenk & Terband, 2020). From a theoretical point of view, this is difficult to explain without further empirical support. Accordingly, on the basis of the DIVA model, a one-to-one ratio of following to opposing responses across a wide range in age is not explicable. Ideas about potentially contributing factors have been provided above. Yet clearly, further research is needed in order to identify the true root of this finding.

6. Conclusion

The comprehensive human aging process encompasses a multitude of changes that occur throughout the lifespan, affecting various aspects of our physical and cognitive abilities. One particular area profoundly influenced by aging is speech, an essential component of oral communication. Previous research has shown a systematic effect of aging on feedback control of the voice pitch (Li et al., 2018; Liu et al., 2011). Furthermore, age-related changes of acoustic and kinematic characteristics of speech production have been demonstrated (Hermes et al., 2018; Tremblay et al., 2018). However, how the human aging process affects auditory-motor adaptation for articulation remained a gap in scientific knowledge thus far. Therefore, this research specifically aimed to explore the impact of aging on auditory-motor adaptation of feedforward control in the articulatory domain.

Based on the results of an auditory feedback shifting experiment focusing on the first formant (F_1), no effect of aging on auditory-motor adaptation for articulation could be identified. However, the possibility for age to influence speech motor control was not ruled out by this outcome. Rather, different subsystems (i.e., feedback and feedforward control) alongside different speech domains (i.e., voice and articulation) might follow separate trajectories of aging. Moreover, there was no interaction between age and sex with respect to adaption magnitudes or trajectories. A surprising finding concerned the direction of adaptive responses rather than the magnitude. Roughly half of the participants opposed the feedback perturbation in their productions, while the other half followed the feedback perturbation, a distinction which has not been reported to this extent in previous work.

In summary, this research emphasized the diversity of aging processes within the specific system of speech motor control. It raised the question of how varied age-related dynamics affecting the subsystems of speech are. Moreover, further research is needed to clarify the origin of the observed distinction in response directions. Insights into factors contributing to this outcome could lead towards novel paths for research on speech motor control.

7. References

Abur, D., Subaciute, A., Daliri, A., Lester-Smith, R. A., Lupiani, A. A., Cilento, D., Enos, N. M., Weerathunge, H. R., Tardif, M. C., & Stepp, C. E. (2021). Feedback and feedforward auditory-motor processes for voice and articulation in Parkinson's Disease. *Journal of Speech, Language, and Hearing Research*, *64*(12), 4682–4694. https://doi.org/10.1044/2021_JSLHR-21-00153

Behroozmand, R., Liu, H., & Larson, C. R. (2011). Time-dependent neural processing of auditory feedback during voice pitch error detection. *Journal of Cognitive Neuroscience 23*(5), 1205–1217. https://doi.org/10.1162/jocn.2010.21447

Bilodeau-Mercure, M., & Tremblay, P. (2016). Age differences in sequential speech production: Articulatory and physiological factors. *Journal of the American Geriatrics Society*, *64*(11), e177–e182. https://doi.org/10.1111/jgs.14491

Boersma, P., & Weenink, D. (2023). *Praat: Doing phonetics by computer* (Version 6.3.06) [Computer software]. http://www.praat.org/

Braun, A., & Friebis, S. (2009). Phonetic cues to speaker age: A longitudinal study. In G. Grewendorf & M. Rathert (Eds.), *Formal Linguistics and Law* (pp. 141–162). Mouton deGruyter.

Buckner, R. L., Andrews-Hanna, J. R., & Schacter, D. L. (2008). The brain's default network: Anatomy, function, and relevance to disease. *Annals of the New York Academy of Sciences*, *1124*(1), 1–38. https://doi.org/10.1196/annals.1440.011

Burnett, T. A., Senner, J. E., & Larson, C. R. (1997). Voice Fo responses to pitch-shifted auditory feedback: A preliminary study. *Journal of Voice*, *11*(2), 202–211. https://doi.org/10.1016/S0892-1997(97)80079-3

Cai, S., Boucek, M., Ghosh, S. S., Guenther, F. H., & Perkell, J. S. (2008, December 8–12). *A system for online dynamic perturbation of formant trajectories and results from perturbations of the Mandarin triphthong /iau/* [Paper presentation]. 8th International Seminar on Speech Production, Strasbourg, France. https://issp2008.loria.fr/

Cai, S., Ghosh, S. S., Guenther, F. H., & Perkell, J. S. (2010). Adaptive auditory feedback control of the production of formant trajectories in the Mandarin triphthong /iau/ and its pattern of generalization. *Journal of the Acoustical Society of America*, *128*(4), 2033–2048. https://doi.org/10.1121/1.3479539

Cai, X., Yin, Y., & Zhang, Q. (2023). Behavioral and ERP evidence of differences in pitch feedback control in late bilinguals' L1 and L2 speech production. *Bilingualism: Language and Cognition*, *First View*, 1–19. https://doi.org/10.1017/s1366728923000019

Cholin, J., & Levelt, W. J. M. (2009). Effects of syllable preparation and syllable frequency in speech production: Further evidence for syllabic units at a post-lexical level. *Language and Cognitive Processes*, *24*(5), 662–684. https://doi.org/10.1080/01690960802348852

Crocco, C. (2017). Everyone has an accent: Standard Italian and regional pronunciation. In M. Cerruti, C. Crocco, & S. Marzo (Eds.), *Towards a new standard: Theoretical and empirical studies on the restandardization of Italian* (pp. 89-117). De Gruyter Mouton.

Daliri, A., & Dittman, J. (2019). Successful auditory motor adaptation requires task-relevant auditory errors. *Journal of Neurophysiology*, *122*(2), 552–562. https://doi.org/10.1152/jn.00662.2018

Esposito, A. (2002). On vowel height and consonantal voicing effects: Data from Italian. *Phonetica*, *59*(4), 197–231. https://doi.org/10.1159/000068347

Eurostat. (2023, May 16). *The share of elderly people continues to increase*. Eurostat. Retrieved May 19, 2023, from https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Population_structure_and_ageing#The_share_of_elderly_people_continues_to_increase

Franken, M. K., Acheson, D. J., McQueen, J. M., Eisner, F., & Hagoort, P. (2017). Individual variability as a window on production-perception interactions in speech motor control. *Journal of the Acoustical Society of America*, *142*(4), 2007–2018. https://doi.org/10.1121/1.5006899

Franken, M. K., Acheson, D. J., McQueen, J. M., Hagoort, P., & Eisner, F. (2019). Consistency influences altered auditory feedback processing. *Quarterly Journal of Experimental Psychology*, *72*(10), 2371–2379. https://doi.org/10.1177/1747021819838939

Grice, M., Savino, M., Caffò, A., & Roettger, T. B. (2015, August 10–14). *The tune drives the text: Schwa in consonant-final loan words in Italian* [Paper presentation]. 18th International Congress of Phonetic Sciences, Glasgow, UK. http://www.internationalphoneticassociation.org/icphs-proceedings/ICPhS2015/proceedings.html

Guenther, F. H. (2006). Cortical interactions underlying the production of speech sounds. *Journal of Communication Disorders*, *39*(5), 350–365. https://doi.org/10.1016/j.jcomdis.2006.06.013

Guenther, F. H. (2016). *Neural control of speech*. The MIT Press. https://www.jstor.org/stable/j.ctt1g69zfb.6

Guenther, F. H., & Vladusich, T. (2012). A neural theory of speech acquisition and production. *Journal of Neurolinguistics*, *25*(5), 408–422. https://doi.org/10.1016/j.jneuroling.2009.08.006

Heinks-Maldonado, T. H., Nagarajan, S. S., & Houde, J. F. (2006). Magnetoencephalographic evidence for a precise forward model in speech production. *Neuroreport*, *17*(13), 1375–1379. https://doi.org/10.1097/01.wnr.0000233102.43526.e9

Hermes, A., Mertens, J., & Mücke, D. (2018, September 2–6). *Age-related effects on sensorimotor control of speech production* [Paper presentation]. Interspeech 2018, Hyderabad, India. https://www.interspeech2018.org/

Hickok, G., & Poeppel, D. (2004). Dorsal and ventral streams: A framework for understanding aspects of the functional anatomy of language. *Cognition*, *92*(1-2), 67–99. https://doi.org/10.1016/j.cognition.2003.10.011

Houde, J. F., & Jordan, M. I. (1998). Sensorimotor adaptation in speech production. *Science*, *279*(5354), 1213–1216. https://www.jstor.org/stable/2894820?seq=1&cid=pdfreference#references_tab_contents

Houde, J. F., & Nagarajan, S. S. (2011). Speech production as state feedback control. *Frontiers in Human Neuroscience*, *5*, 82. https://doi.org/10.3389/fnhum.2011.00082

Hu, H., Li, J., He, S., Zhao, Y., Liu, P., & Liu, H. (2023). Aging-related decline in the neuromotor control of speech production: Current and future. *Frontiers in Aging Neuroscience*, *15*, 1172277. https://doi.org/10.3389/fnagi.2023.1172277

Jiang, J., Lin, E., & Hanson, D. G. (2000). Vocal fold physiology. *Otolaryngologic Clinics of North America*, *33*(4), 699–718. https://doi.org/10.1016/s0030-6665(05)70238-3

Karlin, R., Naber, C., & Parrell, B. (2021). Auditory feedback is used for adaptation and compensation in speech timing. *Journal of Speech, Language, and Hearing Research*, *64*(9), 3361–3381. https://doi.org/10.1044/2021_JSLHR-21-00021

Kearney, E., & Guenther, F. H. (2019). Articulating: The neural mechanisms of speech production. *Language, Cognition and Neuroscience*, *34*(9), 1214–1229. https://doi.org/10.1080/23273798.2019.1589541

Kearney, E., Nieto-Castanon, A., Weerathunge, H. R., Falsini, R., Daliri, A., Abur, D., Ballard, K. J., Chang, S. E., Chao, S. C., Heller Murray, E. S., Scott, T. L., & Guenther, F. H. (2020). A simple 3-parameter model for examining adaptation in speech and voice production. *Frontiers in Psychology*, *10*, 2995. https://doi.org/10.3389/fpsyg.2019.02995

Ketcham, C. J., & Stelmach, G. E. (2004). Movement control in the older adult. In R. W. Pew & S. B. Van Hemel (Eds.), *Technology for adaptive aging* (pp. 64–92). The National Academies Press.

Koenig, L. L., Lucero, J. C., & Perlman, E. (2008). Speech production variability in fricatives of children and adults: Results of functional data analysis. *Journal of the Acoustical Society of America*, *124*(5), 3158–3170. https://doi.org/10.1121/1.2981639

Kothare, H., Raharjo, I., Ramanarayanan, V., Ranasinghe, K., Parrell, B., Johnson, K., Houde, J. F., & Nagarajan, S. S. (2020). Sensorimotor adaptation of speech depends on the direction of auditory feedback alteration. *Journal of the Acoustical Society of America*, *148*(6), 3682–3697. https://doi.org/10.1121/10.0002876

Lametti, D. R., Nasir, S. M., & Ostry, D. J. (2012). Sensory preference in speech production revealed by simultaneous alteration of auditory and somatosensory feedback. *Journal of Neuroscience*, *32*(27), 9351–9358. https://doi.org/10.1523/JNEUROSCI.0404-12.2012

Lametti, D. R., Smith, H. J., Watkins, K. E., & Shiller, D. M. (2018). Robust sensorimotor learning during variable sentence-level speech. *Current Biology*, *28*(19), 3106–3113. https://doi.org/10.1016/j.cub.2018.07.030

Lamster, I. B., Asadourian, L., Del Carmen, T., & Friedman, P. K. (2016). The aging mouth: Differentiating normal aging from disease. *Periodontology 2000*, *72*(1), 96–107. https://doi.org/10.1111/prd.12131

Lester-Smith, R. A., Daliri, A., Enos, N., Abur, D., Lupiani, A. A., Letcher, S., & Stepp, C. E. (2020). The relation of articulatory and vocal auditory-motor control in typical speakers. *Journal of Speech, Language, and Hearing Research*, *63*(11), 3628–3642. https://doi.org/10.1044/2020_JSLHR-20-00192

Levelt, W. J. M. (1989). *Speaking: From intention to articulation*. The MIT Press.

Li, J., Hu, H., Chen, N., Jones, J. A., Wu, D., Liu, P., & Liu, H. (2018). Aging and sex influence cortical auditory-motor integration for speech control. *Frontiers in Neuroscience*, *12*, 749. https://doi.org/10.3389/fnins.2018.00749

Li, Z., & Wood, S. N. (2020). Faster model matrix crossproducts for large generalized linear models with discretized covariates. *Statistics and Computing*, *30*(1), 19–25. https://doi.org/10.1007/s11222-019-09864-2

Liberman, A. M., Cooper, F. S., Shankweiler, D. P., & Studdert-Kennedy, M. (1967). Perception of the speech code. *Psychological Review*, *74*(6), 431–461. https://doi.org/https://doi.org/10.1037/h0020279

Liu, P., Chen, Z., Jones, J. A., Huang, D., & Liu, H. (2011). Auditory feedback control of vocal pitch during sustained vocalization: A cross-sectional study of adult aging. *PLoS One*, *6*(7), e22791. https://doi.org/10.1371/journal.pone.0022791

Markiewicz, C. J., & Bohland, J. W. (2016). Mapping the cortical representation of speech sounds in a syllable repetition task. *Neuroimage*, *141*, 174–190. https://doi.org/10.1016/j.neuroimage.2016.07.023

Matilainen, L. E., Talvitie, S. S., Pekkonen, E., Alku, P., May, P. J., & Tiitinen, H. (2010). The effects of healthy aging on auditory processing in humans as indexed by transient brain responses. *Clinical Neurophysiology*, *121*(6), 902–911. https://doi.org/10.1016/j.clinph.2010.01.007

McFarland, D. H. (2015). *Netter's atlas of anatomy for speech, swallowing, and hearing* (2 ed.). Elsevier.

Miatto, V. (2020, October 12–14). *Perception of word-final inserted vowels and syllabicity in Italian* [Paper presentation]. 11th International Conference of Experimental Linguistics, Athens, Greece. https://doi.org/10.36505/ExLing-2020/11/0034/000449

Mitsuya, T., MacDonald, E. N., Munhall, K. G., & Purcell, D. W. (2015). Formant compensation for auditory feedback with English vowels. *Journal of the Acoustical Society of America*, *138*(1), 413–424. https://doi.org/10.1121/1.4923154

Mücke, D., Thies, T., Mertens, J., & Hermes, A. (2020, December 14–17). *Age-related effects of prosodic prominence in vowel articulation* [Paper presentation]. 12th International Seminar on Speech Production, virtual conference. https://issp2020.yale.edu/

Nair, K. S. (2005). Aging muscle. *The American Journal of Clinical Nutrition*, *81*(5), 953–963. https://doi.org/10.1093/ajcn/81.5.953

Nasir, S. M., & Ostry, D. J. (2009). Auditory plasticity and speech motor learning. *Proceedings of the National Academy of Sciences*, *106*(48), 20470–20475 https://doi.org/10.1073/pnas.0907032106

Niziolek, C. A., & Guenther, F. H. (2013). Vowel category boundaries enhance cortical and behavioral responses to speech feedback alterations. *Journal of Neuroscience*, *33*(29), 12090–12098. https://doi.org/10.1523/JNEUROSCI.1008-13.2013

Oschkinat, M., & Hoole, P. (2020). Compensation to real-time temporal auditory feedback perturbation depends on syllable position. *Journal of the Acoustical Society of America*, *148*(3), 1478–1495. https://doi.org/10.1121/10.0001765

Panouilleres, M. T. N., & Mottonen, R. (2018). Decline of auditory-motor speech processing in older adults with hearing loss. *Neurobiology of Aging*, *72*, 89–97. https://doi.org/10.1016/j.neurobiolaging.2018.07.013

Parrell, B., Lammert, A. C., Ciccarelli, G., & Quatieri, T. F. (2019). Current models of speech motor control: A control-theoretic overview of architectures and properties. *Journal of the Acoustical Society of America*, *145*(3), 1456–1481. https://doi.org/10.1121/1.5092807

Parrell, B., Ramanarayanan, V., Nagarajan, S., & Houde, J. (2019). The FACTS model of speech motor control: Fusing state estimation and task-based control. *PLoS Computational Biology*, *15*(9), e1007321. https://doi.org/10.1371/journal.pcbi.1007321

Perkell, J. S., Guenther, F. H., Lane, H., Matthies, M. L., Stockmann, E., Tiede, M., & Zandipour, M. (2004). The distinctness of speakers' productions of vowel contrasts is related to their discrimination of the contrasts. *Journal of the Acoustical Society of America*, *116*(4), 2338–2344. https://doi.org/10.1121/1.1787524

R Core Team. (2022). *R: A language and environment for statistical computing* (Version 4.2.2) R Foundation for Statistical Computing. www.R-project.org/

Ranasinghe, K. G., Gill, J. S., Kothare, H., Beagle, A. J., Mizuiri, D., Honma, S. M., Gorno-Tempini, M. L., Miller, B. L., Vossel, K. A., Nagarajan, S. S., & Houde, J. F. (2017). Abnormal vocal behavior predicts executive and memory deficits in Alzheimer's disease. *Neurobiology of Aging*, *52*, 71–80. https://doi.org/10.1016/j.neurobiolaging.2016.12.020

Repetti, L. (2012). Consonant-Final Loanwords and Epenthetic Vowels in Italian. *Catalan Journal of Linguistics* *11*, 167–188. https://raco.cat/index.php/CatalanJournal/article/view/292410

Rogers, D., & d'Arcangeli, L. (2004). Italian. *Journal of the International Phonetic Association*, *34*(1), 117–121. https://doi.org/10.1017/s0025100304001628

Scheerer, N. E., Jacobson, D. S., & Jones, J. A. (2016). Sensorimotor learning in children and adults: Exposure to frequency-altered auditory feedback during speech production. *Neuroscience*, *314*, 106–115. https://doi.org/10.1016/j.neuroscience.2015.11.037

Schow, R. L. (1991). Considerations in selecting and validating an adult/elderly hearing screening protocol. *Ear and Hearing*, *12*(5), 337–348. https://doi.org/10.1097/00003446-199110000-00006

Slade, K., Plack, C. J., & Nuttall, H. E. (2020). The effects of age-related hearing loss on the brain and cognitive function. *Trends in Neurosciences*, *43*(10), 810–821. https://doi.org/10.1016/j.tins.2020.07.005

Terband, H., van Brenk, F., & van Doornik-van der Zee, A. (2014). Auditory feedback perturbation in children with developmental speech sound disorders. *Journal of Communication Disorders*, *51*, 64–77. https://doi.org/10.1016/j.jcomdis.2014.06.009

Titze, I. R. (1994). *Principles of voice production*. Prentice Hall.

Tokuda, I. (2021, November 29). The source–filter theory of speech. In M. Aronoff (Ed.), *Oxford Research Encyclopedia of Linguistics*. Oxford University Press. https://doi.org/10.1093/acrefore/9780199384655.013.894

Torre, P., & Barlow, J. A. (2009). Age-related changes in acoustic characteristics of adult speech. *Journal of Communication Disorders*, *42*(5), 324–333. https://doi.org/10.1016/j.jcomdis.2009.03.001

Tourville, J. A., & Guenther, F. H. (2011). The DIVA model: A neural theory of speech acquisition and production. *Language & Cognitive Processes*, *26*(7), 952–981. https://doi.org/10.1080/01690960903498424

Traub-Eichhorn, J., 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), e1–e9. https://doi.org/10.1016/j.jvoice.2017.08.003

Tremblay, K. L., Piskosz, M., & Souza, P. (2003). Effects of age and age-related hearing loss on the neural representation of speech cues. *Clinical Neurophysiology*, *114*(7), 1332–1343. https://doi.org/10.1016/s1388-2457(03)00114-7

Tremblay, P., Deschamps, I., Bedard, P., Tessier, M. H., Carrier, M., & Thibeault, M. (2018). Aging of speech production, from articulatory accuracy to motor timing. *Psychology and Aging*, *33*(7), 1022–1034. https://doi.org/10.1037/pag0000306

Tremblay, P., Sato, M., & Deschamps, I. (2017). Age differences in the motor control of speech: An fMRI study of healthy aging. *Human Brain Mapping*, *38*(5), 2751–2771. https://doi.org/10.1002/hbm.23558

Unell, L., Johansson, A., Ekback, G., Ordell, S., & Carlsson, G. E. (2012). Prevalence of troublesome symptoms related to temporomandibular disorders and awareness of bruxism in 65- and 75-year-old subjects. *Gerodontology*, *29*(2), e772–e779. https://doi.org/10.1111/j.1741-2358.2011.00558.x

Van Brenk, F., & Terband, H. (2020). Compensatory and adaptive responses to real-time formant shifts in adults and children. *Journal of the Acoustical Society of America*, *147*(4), 2261. https://doi.org/10.1121/10.0001018

Van Rij, J., Wieling, M., Baayen, H. R., & Van Rijn, H. (2022). *itsadug: Interpreting time series and autocorrelated data using GAMMs* (Version 2.4.1) [R package]. https://cran.r-project.org/package=itsadug

Villacorta, V. M., Perkell, J. S., & Guenther, F. H. (2007). Sensorimotor adaptation to feedback perturbations of vowel acoustics and its relation to perception. *Journal of the Acoustical Society of America*, *122*(4), 2306–2319. https://doi.org/10.1121/1.2773966

Watson, P. J., & Munson, B. (2007). *A comparison of vowel acoustics between older and younger adults* [Paper presentation]. 16th International Congress of Phonetic Sciences, Saarbrücken, Germany. http://www.icphs2007.de/

Wickham, H. (2016). *ggplot2: Elegant Graphics for Data Analysis* [R package]. Springer-Verlag. https://ggplot2.tidyverse.org

Wieling, M. (2018). Analyzing dynamic phonetic data using generalized additive mixed modeling: A tutorial focusing on articulatory differences between L1 and L2 speakers of English. *Journal of Phonetics*, *70*, 86–116. https://doi.org/10.1016/j.wocn.2018.03.002

Wood, S. N. (2011). Fast stable restricted maximum likelihood and marginal likelihood estimation of semiparametric generalized linear models. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, *73*(1), 3–36. https://doi.org/10.1111/j.1467-9868.2010.00749.x

Wood, S. N., Li, Z., Shaddick, G., & Augustin, N. H. (2017). Generalized additive models for gigadata: Modeling the U.K. black smoke network daily data. *Journal of the American Statistical Association*, *112*(519), 1199–1210. https://doi.org/10.1080/01621459.2016.1195744

Xue, S. A., & Hao, G. J. (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*(3), 689–701. https://doi.org/10.1044/1092-4388(2003/054)

8. Appendices

8.1. Appendix 1: Individual mean baseline F_1

Age	Sex	F_1	Age	Sex	F_1
16	female	485	17	male	415
16	female	519	17	male	427
17	female	444	17	male	450
17	female	495	23	male	493
18	female	437	25	male	406
18	female	456	25	male	450
20	female	471	26	male	400
22	female	470	26	male	411
22	female	440	27	male	451
23	female	445	27	male	674
23	female	450	27	male	441
24	female	476	28	male	381
25	female	503	30	male	475
25	female	479	30	male	395
26	female	469	32	male	587
27	female	443	32	male	400
29	female	406	35	male	447
29	female	644	36	male	385
30	female	481	38	male	440
31	female	473	39	male	351
32	female	474	40	male	406
33	female	459	41	male	444
35	female	466	41	male	419
36	female	472	44	male	496
37	female	444	50	male	412
37	female	535	56	male	437
38	female	417	57	male	457
38	female	433	58	male	425
39	female	468	60	male	442
42	female	504	62	male	425

43	female	513	62	male	514
45	female	469	65	male	433
48	female	427	70	male	414
49	female	436	75	male	437
51	female	538	82	male	467
54	female	482			
55	female	477			
56	female	450			
58	female	484			
60	female	527			
60	female	566			
63	female	484			
66	female	414			
66	female	504			
72	female	555			

Note. Table complementing Figure 2, showing individual participants' mean raw F_1 values of the baseline trials, F_1 measured as the mean of a time window of 40 to 120 ms within the first /eː/ in consonant-/eː/-consonant-/e/ sequences.

8.2. Appendix 2: Response magnitudes (% F_1 deviation in the hold phase)

Age	Response direction	Norm. F_1	Age	Response direction	Norm. F_1
17	following	8.83	16	opposing	-8.33
20	following	1.91	17	opposing	-2.17
22	following	6.18	17	opposing	-8.46
23	following	6.81	17	opposing	-3.84
23	following	0.10	17	opposing	-6.69
23	following	5.67	18	opposing	-4.36
26	following	3.60	18	opposing	-8.08
26	following	1.70	22	opposing	-1.90
27	following	11.63	24	opposing	-9.00
27	following	2.42	25	opposing	-1.76
27	following	5.25	25	opposing	-12.94
29	following	10.71	25	opposing	-8.17
29	following	2.71	25	opposing	-9.64
30	following	3.95	26	opposing	-4.83

31	following	4.66	27	opposing	-8.02
33	following	3.28	28	opposing	-0.09
35	following	7.86	30	opposing	-7.00
37	following	12.41	30	opposing	-0.54
39	following	10.12	32	opposing	-12.17
39	following	12.29	32	opposing	-3.10
40	following	0.20	32	opposing	-0.30
42	following	9.09	35	opposing	-4.43
43	following	8.45	36	opposing	-3.86
44	following	2.90	36	opposing	-10.89
45	following	1.81	37	opposing	-3.85
48	following	7.92	38	opposing	-0.42
49	following	4.28	38	opposing	-1.19
50	following	2.69	38	opposing	-11.24
55	following	6.89	41	opposing	-2.61
56	following	2.54	41	opposing	-6.31
57	following	10.97	51	opposing	-5.16
58	following	2.39	54	opposing	-2.41
60	following	22.53	56	opposing	-0.98
62	following	0.95	58	opposing	-1.38
62	following	0.98	60	opposing	-2.52
63	following	0.66	60	opposing	-3.56
66	following	0.39	65	opposing	-1.11
73	following	16.08	66	opposing	-8.81
82	following	3.43	70	opposing	-5.87
			72	opposing	-1.40
			75	opposing	-3.26

Note. Table complementing Figure 8, column 'norm. F_1' lists individual participants' response magnitudes (i.e., mean F_1 deviation across the hold trials as compared to the baseline trials).