Speech Communication 141 (2022) 28–39

Contents lists available at ScienceDirect

Speech Communication

journal homepage: www.elsevier.com/locate/specom

Learning transfer from singing to speech: Insights from vowel analyses in aging amateur singers and non-singers

Anna Marczyk [a], Émilie Belley [b,c], Catherine Savard [b,c], Johanna-Pascale Roy [d], Josée Vaillancourt [e], Pascale Tremblay [b,c,*]

[a] Aix-Marseille University, CNRS, LPL, UMR 7309, Aix-en-Provence, France
[b] CERVO Brain Research Center, Quebec City G1J 2G3, Canada
[c] Département de Réadaptation, Université Laval, Faculté de Médecine, 1050 avenue de la Médecine, Office 4109, Quebec City G1V 0A6, Canada
[d] Département de langues, linguistique et traduction, Université Laval, Quebec City G1V 0A6, Canada
[e] Faculté de Musique, Université Laval, Quebec City G1V 0A6, Canada

ARTICLE INFO

Keywords:
Aging
Music
Singing
Vowel production
Speech production
Non-speech

ABSTRACT

Purpose: Task-independent (*e.g.*, Ballard et al., 2003) and task-dependent models (*e.g.*, Ziegler, 2003) differ in their predictions regarding the learning transfer from non-speech activities to speech. We argue that singing is a musical and essentially oral motor activity that differs from speech on several fundamental aspects. Building upon existing evidence on the benefits of vocal training on speech, this study aimed to examine whether amateur singing practice had a protective effect on vowel acoustic identity (speech) in aging speakers.
Method: 75 healthy non-singers (mean age 55.52 ± 20.32; 20–98 years, 39 females) and 71 healthy singers (mean age 55.34 ± 19.34; 21–87 years, 44 females), all native Quebec French speakers, were recruited. Participants were asked to read aloud a standardized passage ("La bise et le soleil"). Vowel duration as well as the F1 and F2 values for the oral vowels /a, i, e, ɛ, o, ɔ, u, y, œ / were extracted from the recordings. A multiparametric assessment of spectral and temporal vowel features was carried out.
Results: The results revealed that spectral characteristics of vowels—such as the size of vowel space and vowel distinctiveness—were overall well preserved across the lifespan, while temporal characteristics—such as speech rate—declined. Singing practice was associated with the size of vowel space in female speakers but not with vowel clarity in either female or male speakers .
Conclusions: Contrary to our hypotheses, amateur singing had little effect on vowel articulation in read speech, likely because vowel quality was not substantially altered by the aging process in this context. Overall, our findings provide support to the task-dependent account of nonspeech-to-speech learning transfer and suggest that complementary analyses of articulatory and linguistic aspects of vowel characteristics may provide relevant insights on potential targets of change.

1. Introduction

Normal aging is associated with disruptions in speech production, affecting the temporal properties of speech, such as articulation rate, articulation rate stability, but also articulation accuracy and, frequently, intelligibility (Kuruvilla-Dugdale et al., 2020). These changes can lead to increased self-consciousness about one's communication competences, reduced self-confidence, and disengagement from social activities, which can lead to isolation (Tobias, 1977). Specifically, aging has been associated with an increase in the duration and duration variability of speech utterances during syllable and sentence repetition (*e.g.* Morris and Brown 1987, Smith et al. 1987), syllable and nonword (*i.e.* sets of syllables not forming real words) reading (Tremblay and Deschamps, 2016; Tremblay et al., 2018, 2017), and nonword repetition (Sadagopan and Smith, 2013). Others have reported a decline in articulation rate in diadochokinetic (DDK) tasks (Bilodeau-Mercure and Tremblay, 2016; Jacewicz et al., 2010; Padovani et al., 2009). In addition to timing-related decline, several studies have reported a decline in

* Corresponding author at: Département de Réadaptation, Université Laval, Faculté de Médecine, 1050 avenue de la Médecine, Office 4109, Quebec City G1V 0A6, Canada.
E-mail address: Pascale.Tremblay@fmed.ulaval.ca (P. Tremblay).

https://doi.org/10.1016/j.specom.2022.05.001
Received 25 July 2021; Received in revised form 15 April 2022; Accepted 5 May 2022
Available online 7 May 2022
0167-6393/

articulation accuracy in nonword repetition (Sadagopan and Smith, 2013), and in syllable, nonword and sentence reading (Bilodeau--Mercure et al., 2015; Gollan and Goldrick, 2018; Tremblay et al., 2018). A recent study suggests that aging is associated with lower tongue movement speed (Kuruvilla-Dugdale et al., 2020). Accuracy decline is worse for long compared to short nonwords, and for phonologically complex syllables compared to simpler ones (Bilodeau-Mercure et al., 2015; Sadagopan and Smith, 2013). Though less frequently studied, vowel space, which can influence intelligibility, has also been shown to vary as a function of age. A recent analysis of the "Up" corpus (Gahl et al., 2014), which is based on a series of documentary films featuring a group of individuals aged between 21 and 49 years, filmed at seven-year intervals, over a period of 56 years (Apted, 1977, 1984, 1991, 1998), reported a shift in vowel space (formed by the first two vowel formants) towards the periphery from young to middle-aged adulthood (Gahl and Baayen, 2019). No data were available on old adults. In sum, speech production undergoes a number of transformations with age, but it remains unclear whether these changes are causing the reduction in intelligibility that is characteristic of elderly speech (Kuruvilla-Dugdale et al., 2020), and whether these changes can be alleviated or even prevented through the practice of targeted activities involving the vocal apparatus.

1.1. Singing and speech intelligibility

One potential activity to maintain speech intelligibility in aging is amateur singing.

Singing is a low-cost universal human activity that belongs to the domain of music but relies on both musical (melody) and verbal (lyrics) components (Patel and Peretz, 1997). Musicality—typically understood as instrumental training—has been shown to facilitate speech processing, from enhancement of phonological processing and sensorimotor speech control mechanisms to increasing perceptual sensitivity to shared basic acoustic features such as duration, pitch and intensity (for a review, see Besson et al. 2011). While scarcer, studies on vocal training effects have reported similar positive influences on speech-related behavior or underlying speech neural circuity. For example, singing proficiency was found to predict speech imitation abilities (Christiner and Reiterer, 2013), which are part of overall speech motor planning skills (Van Der Merwe, 2021). Turning to neuroimaging studies, Kleber et al. (2010) have shown that a long-term vocal training leads to increased activity in the primary somatosensory cortex responsible for proprioceptive feedback and motor accuracy. Halwani et al. (2011) reported increased functional connectivity between motor and auditory systems, which were attributed to singing-induced microstructural changes to the arcuate fasciculus, a major white-matter tract known for its critical role in sensorimotor integration (Dick et al., 2014; Dick and Tremblay, 2012). Interestingly, singers from the Halwani *et al.*'s study (all classically trained singers) showed a significant advantage not only over non-musicians, but also over instrumentalists, suggesting that singing may be particularly advantageous for speech. Taken together, the above findings indicate that musical abilities—whether instrumental or vocal— enhance speech processing mechanisms such as auditory working memory, acoustic and sensory perception and sensorimotor integration, which are essential to speech intelligibility (Luo et al., 2017; Souza et al., 2015). Thus, singing practice as a musical activity may also have beneficial effects on speech intelligibility.

While singing demonstrably belongs to the musical domain, singers differ from instrumentalists in that they use their voice and vocal apparatus as an instrument. Hence, singers intrinsically exhibit greater control over the respiratory system and have greater aerodynamic capacity (Ravi et al., 2019). They also operate on a wider range of intensities and frequencies (Story, 2004) and show greater motor dexterity and motor control. Data from pathological speech such as dysarthria demonstrates that speech intelligibility relies heavily on oral motor ability (Kent et al., 1989). Thus, singing practice as a complex oral motor activity (an oral motor gymnastics so to speak) may also have beneficial effects on speech intelligibility.

Turning to the verbal component of singing practice, song-to-speech benefits are less straightforward. While singing has a clear verbal component and exhibits an intuitive affinity to speech ('Singing and speaking are underpinned by the same speech generation process', Christiner and Reiterer, 2013:8), it deviates from it on a few fundamental elements. The crucial difference—in what concerns the present paper—is that speech seeks to enhance message intelligibility, whereas singing sacrifices it for esthetic purposes. Consequently, acoustic targets for speech intelligibility and singing quality differ in that speech emphasizes linguistic information encoded in lower frequencies, particularly the first three formants (F1, F2, F3), whereas singing enhances sound quality cued in higher frequencies, above F3 (Gregg and Scherer, 2006; Story, 2004). This crucial difference is functional in nature: while both use vocal apparatus to produce acoustic signal, speech and singing target different acoustic frequency zones. Unsurprisingly, intelligibility loss has been frequently reported for both sung consonants and—to a greater extent—vowels, known to carry melodic variations (Collister and Huron, 2008; Deme, 2014; Gregg and Scherer, 2006; Story, 2004). These studies reported negative correlations between high-pitched sung vowels and intelligibility, particularly for closed vowels where the pitch and first vowel formant co-occur around 300 Hz (*e.g.* Hollien et al., 2000; Story, 2004), as well as a tendency towards vowel centralization (Collister and Huron, 2008; Hollien et al., 2000). Another notable difference between spoken *vs*. sung sounds that affects speech intelligibility concerns coarticulation (Benguerel and Ulkrainetz Mcfadden, 1989). Formant transitions, for example, are ubiquitous in speech but radically reduced in singing (Deme, 2014), which can be particularly detrimental to consonant intelligibility cued in these transitions. The dissociation between singing and speech goals outlined above is also supported by clinical data (double dissociation between aphasia and amusia (*e.g.*, Hebert and Peretz, 2001) and neuroimaging studies (*e.g.* Albouy et al., 2020; Ozdemir et al., 2006). This evidence suggests that singing and speaking are relatively independent skills, with singing being more on the oral motor side of the speech – non-speech behavior spectrum. In the current paper, we propose that singing, as a combined musical and oral motor activity, will have a positive influence on articulatory dexterity and consequently, on vowel intelligibility.

1.2. Learning transfer from singing to speech

The notion that singing could impact speaking—that is, the learning transfer from oral motor (non-speech) to oral verbal (speech)—is consistent with the Integrative Model of Speech Motor Control (also referred to as the task-independent account, see Ballard et al., 2003) which proposes that speech and non-speech orofacial functions are controlled, at least in part, through domain-general brain networks, and that working on one behavior (*e.g.*, singing) might have beneficial effects on another (*e.g.*, speaking). Given that, in this perspective, singing and speaking share the same apparatus, which includes the respiratory system, the vocal tract, and the articulators (such as the tongue, soft palate, and lips), an impact of one behavior on the other is expected. The notion of an impact of singing on speech skills is also consistent with the Overlap, Precision, Emotion, Repetition, Attention (OPERA) hypothesis (Patel, 2011, 2012, 2014), which proposes that musical activities affect speech skills by driving plasticity within neural circuitry shared between musical activities and speech processes. Finally, several studies have shown a positive effect of singing on the aging voice, though with heterogeneous benefits (for a review, see Tremblay and Veilleux, 2018). The effect of singing on speech production has been less thoroughly investigated, but a recent meta-analysis showed that a formalized singing-based speech-language intervention, the Melodic Intonation Therapy (MIT, Albert et al., 1973; Sparks et al., 1974), is associated with speech improvements in participants with motor speech disorders (Zumbansen and Tremblay, 2018), supporting the notion that singing *training* can have an impact on speech production, at least in speakers with motor speech disorders.

On the other hand, the hypothesis that a primarily non-speech oral motor activity such as singing could influence speech is incongruous with a more linguistically oriented account (known as the task-dependent model, see Bunton 2008, Weismer 2006, Ziegler, 2002). On a categorical version of this account, speech and non-speech motor functions are two fundamentally different sets of skills supported by specifically dedicated neural networks and controlled by different sensorimotor mechanisms. The lack of shared motor control systems postulated within such a viewpoint does not allow to account easily for a learning transfer from a musical, primarily oral motor activity (singing) to speech (see Maas, 2017 for a similar argument).

Within these opposing frameworks, the general purpose of the study was to determine whether and how a learning transfer from singing to articulation occurs in healthy adult speakers, as measured on vowel acoustic characteristics. To answer the *how* question, two aspects related to the opposing accounts on singing-to-speech learning were considered: vowel articulatory working space and vowel distinctiveness. The first describes the *articulatory potential* for vowel articulation, while the second gives a more precise measure of the *linguistic value* of vocalic realizations, that is, how overlapping—and thus confusing—vowel productions are. These two types of metrics are relatively independent in the sense that a large vowel space does not necessarily ensure a greater distinctiveness and conversely, high distinctiveness can be achieved within a smaller articulatory working space (Huet, 2000; Meunier, 2018). The task-independent account predicts that learning effects—if any—will be observed on both vowel space and vowel distinctiveness, whereas the task-dependent account predicts that singing effect might influence non-linguistic aspects of vowel quality, such as vowel space size, but not vowel distinctiveness. Therefore, if learning effects are observed on vowel distinctiveness, such evidence will support task-independent accounts and speak against the opposite account, at least in its radical version (*i.e.*, without a clear specification of transfer mechanisms).

Turning to age effects, the first hypothesis was that vowel acoustic quality would be lower in older compared to younger adults, a degradation that is typically captured by spectral metrics. Specifically, we expected smaller vowel space and reduced vowel distinctiveness in older compared to younger adults. The second hypothesis was that vowel duration and, more globally, articulation rate would exhibit age-related differences, with longer duration and slower pace in older adults, consistent with previous studies. Additionally, we examined the interaction between spectral (vowel space, vowel distinctiveness) and temporal (duration and articulation rate) aspects of vowels in aging singers and non-singers. Since this interaction has never been probed, its inclusion was largely exploratory. We expected that the pattern of change in aging speakers, characterized above, would be significantly reduced in amateur singers compared to non-singers, reflecting a protective effect of singing on speech functions. It is in this group of speakers that singing-to-speech learning effects were expected to be higher.

2. Method

2.1. Participants

A total of 152 healthy adults (mean age 52.88 ± 20.18; 20–98 years, 87 females) were recruited to participate in this cross-sectional group study through emails, Facebook posts and posters distributed in the general community and at Université Laval, and through emails and Facebook posts targeting all choirs in the Quebec City area. The study was approved by the *Comité d'éthique de la recherche sectoriel en neurosciences et santé mentale, Institut Universitaire en Santé Mentale de Québec* (#192–2017).

All participants were non-smoking native speakers of Québec French, had normal or corrected-to-normal vision and no self-reported speech, language, psychological, neurological, or neurodegenerative disorder. Furthermore, participants reported no past or present diagnosis of voice, swallowing or respiratory disorder, and none had uncontrolled gastric reflux. Participants were screened for depression using the Geriatric Depression Scale (GSD) (Yesavage et al., 1982) and their cognitive level was assessed using the Montreal Cognitive Assessment scale (MoCA) (Nasreddine et al., 2003). From the initial sample, five participants were excluded because they did not respect the inclusion or exclusion criteria.

The final sample (N = 146) was divided into 75 non-singers (mean age 55.52 ± 20.32; 20–98 years, 39 females) and 71 amateur singers (mean age 55.34 ± 19.34; 21–87 years, 44 females). Singers and non-singers did not differ in age, education, handedness, number of spoken languages, MoCA, depression, self-perceived health, or hearing (Table 1). Participants' characteristics are reported in Table 1.

All participants answered a questionnaire on past and present singing and musical habits (Table 2). Choral singers were defined as individuals singing in a choir for at least 2 years and with a weekly choral practice of at least 60 consecutive minutes. Non-singers were defined as individuals not involved in any form of group singing, and not singing professionally. Singers had an average of 16.87 years of singing experience, ranging from 2 to 62 years. Most started singing as adults (63/71), and most never received formal singing training (43/71). In addition to the weekly choir, 40% of all singers practised singing at home every day (28/71) while the same proportion practised at home once a week (28/71) and the rest practised at home either once a month or less (11/71). These singing parameters detail the amount and type of a singer's practice and can affect music-related neuroplasticity (for a review, see Merrett et al., 2013).

2.2. Procedure

The experiment took place in a double-walled sound-attenuated room at the Speech and Hearing Neuroscience Laboratory in Quebec City, Canada.

2.2.1. Audiometric evaluation

Pure-tone thresholds in dB HL were measured with a calibrated clinical audiometer (AC40, Interacoustic, Danemark). The following frequencies: 0.5, 1 and 2 kHz, were assessed in each ear separately following the Hughson-Westlake procedure (Carhart and Jerger, 1959). These measurements were used to compute a better ear (*i.e.*, lowest thresholds) pure tone average (PTA). Singers and non-singers did not differ in better ear PTA (p = .39), or inter-aural difference (p = .45). There were a few participants in both groups with thresholds above 25 dB HL. Because normal age-related hearing impairment in adults can affect *intelligibility*, that is, the proportion of a speaker's output that a listener can readily understand (*e.g.* Perkell et al. 2000, 2007), hearing (better ear PTA) was included as a covariate in all statistical analyses.

2.2.2. Recordings

All recordings were performed under identical conditions in a double-walled sound-attenuated room. Participants were seated in a comfortable armchair. They were provided with a written version of the standardized passage ("La bise et le soleil") on a plastic sheet and asked to first read it quietly . Next, they were asked to read the passage aloud as naturally as possible. They were told that the task served to analyze their oral language production and that they would be recorded. Participants' responses were recorded using a head-worn Shure headset microphone (Microflex Beta 53) connected to a Quartet USB audio interface (Apogee Electronics, Santa Monica, USA) that fed into an iMac computer. The recordings were made using the Sound Studio 4 software (Felt Tip Inc., NYC, USA) at a sampling signal of 48 kHz and 24 bits of quantization.

The reading task was part of a larger project in which voice and speech functions were assessed including measures of sustained vowel production, maximal phonation, pitch and intensity voice crescendo,

Table 1
Participant's characteristics separately for each group.

Characteristic	Non-singers (N = 75, F = 39) Mean	SD	Min	Max	Singers (N = 71, F = 44) Mean	SD	Min	Max	Group difference t	p
Age	52.52	20.32	20.00	98.00	55.34	19.34	21.00	87.00	−0.86	0.39
Education in years	15.11	2.79	9.00	21.00	14.87	2.52	6.00	23.00	0.89	0.38
Handedness	0.11	0.31	0.00	1.00	0.07	0.31	0.00	2.00	0.71	0.48
Nb languages	2.25	1.00	1.00	7.00	2.20	0.60	1.00	4.00	0.41	0.68
MoCA(/30)	27.37	2.15	21.00	30.00	27.08	2.71	17.00	31.00	0.72	0.48
GDS (/30)	2.84	2.99	0.00	14.00	2.45	3.21	0.00	19.00	0.76	0.45
Health (/7)	5.26	0.89	3.00	7.00	5.02	1.10	1.00	7.00	1.45	0.15
Better ear PTA	10.13	9.91	−5.00	38.33	8.80	8.61	−6.67	30.00	0.87	0.39

Note. N = number of participants; F = number of female participants; M = Mean; SD = standard deviation;.
[a]Education = Number of years of education based on the highest degree obtained.
[b]Handedness = The handedness was measured with the Edinburgh Handedness Inventory. Participants report which hand they use to perform ten actions on a scale: always the same hand (2 points), usually the same hand (1 point), without preference (0 point). Based on results a lateralization quotient is calculated: 100 * (score for the right hand - score for the left hand)/ 20. A quotient of 60% or more indicates laterality on the right.
[c]Nb languages = number of spoken languages including the native language (French).
[d]MoCA = Montreal Cognitive Assessment scale. The MOCA is a short cognitive test that is scored on a 30-point scale. Higher scores indicate better cognitive functions.
[e]GDS = Geriatric Depression Scale. The GDS includes 30 yes/no questions. Each "negative" answer is worth one point; thus, a higher score indicates a more depressed state. For example, question one asks whether the person is globally satisfied with his/her life. A "no' answer is worth one point, whereas a "yes" answer is worth no point. Participants with scores between 0 and 9 are considered normal, while scores between 10 and 19 indicate a depression, and scores between 20 and 30 indicate a severe depression.
[f]Health = self-reported general health status on a scale of 0 to 7, with 0 being lowest health level and 7 the maximal one.
[g]Better ear PTA = pure tone average thresholds (PTA) at 0.5, 1, 2, 4 and 6 kHz for the better ear, measured in decibels (dB).

Table 2
Singers' characteristics.

	Mean	SD	Min	Max
Age at onset (years)	38.46	18.49	7	84.5
Singing experience (years)	16.87	14.31	2	62
Singing ratio (singing experience/age) (%)	30.77	22.16	2.87	79.41

DDK, nonword repetition and storytelling. These measures are not reported here. A speech perception in noise task was also administered and MRI data were acquired. The speech perception and some of the MRI data analyses are reported in Perron et al (2021, 2022).

2.2.3. Transcription and segmentation

The data were analyzed with Praat (Boersma and Weenink, 2011). First, the recordings were automatically aligned and segmented using the EasyAlign Praat plugin (http://latlcui.unige.ch/phonetique/easyalign.php). A macro-segmentation at the utterance level was first performed, which was followed by a grapheme-to-phoneme conversion and a final phone segmentation. The automated segmentation was submitted to quality control: three independent evaluators corrected the automated alignment manually when needed (EB, CS, AM). Intraclass correlation coefficients (ICC) between the judges and their 95% confident intervals were calculated using SPSS V 26 (SPSS Inc, Chicago, IL) based on a mean-rating, absolute-agreement, 2-way mixed-effect model. A high degree of reliability was found among the judges indicating high consistency in segmentation. The single measure ICC was 0.931 with a 95% confidence interval from 0.921 to 0.940 ($F(447,1788) = 69.28$, $p < .0001$).

Next, a Praat script was used to extract vowel duration, as well as the F1 and F2 values for the following oral vowels: /a, i, e, ɛ, ə, o, ɔ, u, y, œ/ (representing a total of 20 312 tokens). To avoid possible coarticulation effects, mean formant values were extracted from the middle third of each vocalic segment (the most steady-state vowel segment). Vocalic segments shorter than 30 ms and longer than 400 ms (likely corresponding to filled pauses) were excluded from further analyses (excluded segments represented 0.65% of the total vowels). A filter (Gendrot and Adda.-Decker, 2005)was applied to detect and correct aberrant values manually. The filter provides upper and lower possible formant values for French male and female speakers. Any value outside these boundaries was removed (1.1% of all formant measurements).

2.2.4. Vowel quality metrics

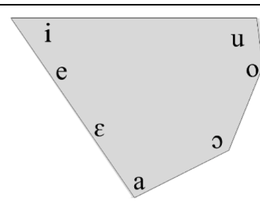
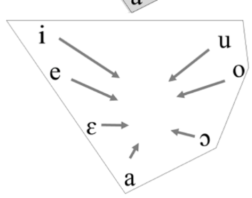
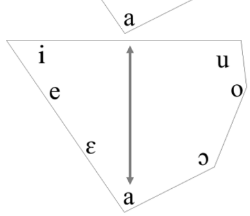
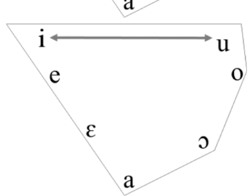
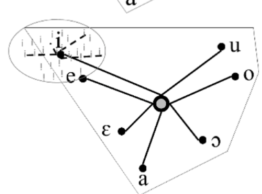
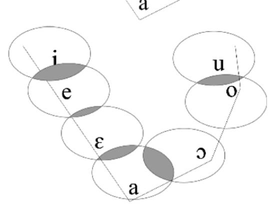
Because aging can affect both spectral and temporal characteristics of vowel production, a twofold vowel quality assessment was carried out. Spectral measurements provide information about the range of variation in token realization, that is, speakers' ability to reach phonemic targets. Since greater variability may reduce vowel distinctiveness and affect the listener's perception, spectral assessment is ultimately informative about how intelligible and comprehensible one's speech is, and thus, is related to the notion of speech intelligibility. Spectral measurements are extracted from a two-dimensional F1/F2 plane and aim at quantifying the shifting and/or contraction *vs.* expansion of the vowel system as a function of the predictor variables. If, by our main hypothesis, singing practice has a protective effect on vowel clarity and can reduce—to a certain degree—vowel centralization processes associated with age, the spectral metrics should account for a gradient nature of changes we expected to observe. Indeed, several previous studies on vowel quality in speech impairments or sociophonetics have suggested that some acoustic metrics may be too global to capture small alterations in vowel quality (see the next paragraphs). Hence, we opted for a multiparametric spectral assessment using several complementary metrics that rely on different computational principles and capture different aspects of the vowel production dynamics, specifically, the parameters of the articulatory working space and vowel distinctiveness.

Metrics of vowel articulatory working space. The Pentagonal Vowel Space Area (pVSA), one of the most common metrics of vowel quality, is expressed as the total articulatory area defined by the peripheral vowels, which in French has the geometric form of a polygon limited by the corner vowels /i, y, u, e, ɛ, ɔ, o, a/. The pVSA metric provides a good measure of articulatory working space and is a fairly reliable estimate of vocal tract lengthening (simultaneous decrease in all formant frequencies that results in shifting of the vowel space towards the upper right corner of the F1/F2 plane (Laver, 1991). It has been shown to be fairly sensitive to robust factors of variation such as sex or speaking style—with smaller areas being associated with articulatory reduction—, but proved less performant in capturing more subtle changes as those in milder forms of speech impairments (Sapir et al., 2010; Skodda et al., 2012) or within-dialect regional differences (Fox, 2017). In contrast, acoustic metrics expressed as a ratio allows for the registration of small alterations in vowel quality, such as varying degrees of severity or treatment effects in clinical research (Sapir et al., 2010), while masking effects due to inter-speaker variability such as sex or age and thus seem appropriate for our purposes. Four such metrics were retained for the study. An adaptation of the Formant Centralization Ratio (aFCR Sapir et al., 2010), provides an estimate of vowel compression by comparing formant frequencies likely to increase (the numerator) with those likely to decrease (the denominator) with vowel centralization (for its application to French see Audibert, 2012; Martel Sauvageau et al., 2014). The aFCR values increase with system compression and decrease with system expansion and have been shown to be optimal for registering subtle changes induced by a speech pathology, while being minimally affected by inter-speaker variability. Finally, we used the First Formant Range Ratio (F1RR) and the Second Formant Range Ratio (F2RR), designed to measure lingual mobility roughly corresponding to tongue height (F1) and tongue body advancement (F2), to assess more closely reduction processes associated with either of these dimensions.

Vowel distinctiveness metrics. A composite Vowel Distinctiveness Index (VDI; see Huet 2000, Meunier 2018) was used to estimate distances between vowel categories and the whole vowel system that allows to account for compensatory dynamics that may likely take place between these two elements. VDI relies on the relationship between the total vowel space and dispersion of each vowel category. Vowel and system dispersion were defined in terms of mean Euclidean distances between the vowel/system centres and their periphery. The index was computed following the approach described in Meunier and Ghio (2018). Higher VDI scores are associated with greater distinctiveness and better speech intelligibility. Both Mean Vowel System Dispersion (mVSD) and Mean Vowel Category Dispersion (mVCD) were also saved for the analyses.

If reduction processes can take the form of vowel centralization (all vowels sound like a schwa and occupy the center of the vowel system), the contrast loss may also occur by merging vowel categories that lie close to each other within the vowel area. To quantify and assess such processes, we measured the degree of overlap between neighbouring vowel categories by means of the Pillai-Bartlett Trace (henceforth Pillai scores) for vowel pairs /i, e – e, ε/ - ε, a – a, ɔ - ɔ, o – o, u/ in the F1 and F2 dimension (Hay et al., 2006; Kelley and Tucker, 2020; Nycz and Hall--Lew, 2014). The Pillai score is a summary statistic of a multivariate analysis of variance (MANOVA) that reflects the degree of separability between two distributions. Previous studies have shown that overlap metrics not only significantly distinguished between speakers with dysarthria and healthy speakers (*i.e.*, registered subtle acoustic changes (*e.g.* Audibert 2012), but also accounted for the greatest amount of variance when correlated with perceptual metrics of intelligibility (Kim

Table 3
Summary of spectral metrics of vowel quality, their schematic representation and computation formulae.

Metric	Schematic representation	Formula
Vowel space area (pVSA) (Chung et al., 2012; Kent and Kim, 2003; Vorperian and Kent, 2007)		*vowelMeansPolygonArea* function in phonR package, (McCloy, 2016). Unlike in previous studies, all the bordering vowel categories depicted on the left were considered for the vowel space area calculation.
Formant centralization ratio (aFCR) (Audibert, 2012; Sapir et al., 2010)		$\frac{F2u + F1i + F1u + F2o + F2\,(increasing\,formant\,values)}{F2i + F1a + F2e + F2(decreasing\,formant\,values)}$ The arrows reflect the increasing/decreasing F1 and F2 values
First formant range ratio (F1RR) (Sapir et al., 2010)		$\frac{2\ F1a}{F1i + F1u}$
Second formant range ratio (F2RR) (Sapir et al., 2010)		$\frac{F2i}{F2u}$
Vowel Distinctiveness Index (VDI) and its composites **Mean Vowel System Dispersion** (mVSD) and **Mean Vowel Category Dispersion** (mVCD) (Huet, 2000; Meunier, 2018)		$\frac{\bar{x}\ Vowel\,system\,dispersion\ (solid\,lines)}{\bar{x}\ Vowel\,category\,dispersion\ (dashed\,lines)}$
Pillai-Bartlett Trace (Pillai scores) (Hay et al., 2006; Nycz and Hall-Lew, 2014)		MANOVA statistic summary

et al. 2012). Pillai scores range from 0 to 1 with higher values indicating greater separability and those close to 0 overlapping distributions. After obtaining values per each vowel pair and participant, a mean Pillai score per participant was computed. The summary of all the spectral metrics used in the study including their mathematical expressions is provided in Table 3. The detail of the relationship between these different metrics is provided in Supplementary Material 2.

2.2.5. Temporal assessment

Two temporal metrics were examined: Articulation Rate and Mean Vowel Duration.

Articulation rate. Articulation rate was defined as the pace at which speech segments were produced by a speaker without considering silent pause intervals. We chose articulation rate over alternative pace metrics such as speech rate (typically using syllables per second as a proxy measure), because it is less sensitive to pragmatic phenomena such as pauses, hesitations or fillers, and provides a less global and more accurate estimate of the actual articulation rate (rather than the prosodic, rhythmic aspect of speech). Articulation rate was defined operationally as the number of segments (all vowels and consonants) per second and was extracted automatically for each participant by means of the EasyAlign Praat plugin (see § 2.2.3 for the detail of the segmentation process and its verification).

Vowel Duration. Mean Vowel Duration was obtained by averaging the duration of all vowels per speaker. It shall be noted that vowel durations were extracted based on the forced alignment of speech samples by EasyAlign (see § 2.2.3), which uses the acoustic information (abrupt changes in energy, F2/F3 onset and offset) to align the signal to the provided phonetic content. Thus, contrary to spectral analyses which focused on the stable portion of the vowel (the middle third), for vocalic duration analyses formant transitions were included.

Articulation rate and mean vowel duration were inversely correlated (r=−0.79, p=.000). Considering that all speakers read the same passage, we assumed that vowel type, syllabic structure, stress and prosodic patterns influenced all speakers in the same way and did not include these variables in our analytical framework. All computations were performed using customized R and Praat scripts.

2.2.6. Statistical data analyses

We used the Kolmogorov-Smirnov test to assess the normality of all the dependent variables in each sex group. Since all the tests indicated Gaussian distributions, parametric statistics (multivariate linear regression) were used for all the analyses. Apart from the model statistics, we opted for using type I test (ANOVA) to assess the significance of main effects and interactions instead of commonly used t-statistics (type III), as it provides more powerful and more easily interpretable tests of the hypotheses, particularly for factors (such as singing practice) and interactions, which are the focus of this study. Post hoc comparisons for categorical variables were performed using the E*mmeans* package in Lenth et al. (2020), R core Team, 2019 and standardized beta coefficients (measured in units of standard deviation) were reported for continuous variables to allow comparisons in terms of impact of predictor variables encoded on different scales. The significance level was set at p<.05 with no adjustments for multiple comparisons as the study focused on planned comparisons that were part of the experimental design (Rothman, 1990; Saville, 1990). Standardized mean difference (Cohen's d) effect sizes were calculated for all the significant effects. As mentioned earlier, vocal tract dimension differences in men and women are associated with robust phonetic effects. Because such effects could mask more subtle acoustic changes attributable to the primary variables of interest (that is, age and singing practice), hypothesis testing was performed separately for male and female speakers , depending on the outcome of an initial analysis of sex effects on spectral and temporal variables. For hypothesis testing, age (continuous variable), singing practice (singers *vs.* non-singers) and an interaction between these two factors were entered as main predictors. It is well established that auditory acuity and speech tempo influence vowel quality, thus, to account for its possible effects, a metric of auditory acuity (better ear PTA) and temporal variables (Articulation Rate and Mean Vowel Duration) were entered as moderating variables in each model. We used the backward elimination procedure in model optimization by excluding variables with a p-value of >0.05 in a stepwise manner (Hong and Mitchell, 2007). Data visualization was carried out using the *ggplot2* (Wickham, 2016) and *phonR* (McCloy, 2016) packages in R.

3. Results

3.1. Spectral analyses

The analyses for male speakers revealed that the model based on Pillai score were more performant as compared to all the other single metric-based models, explaining 16% of the variance. For the sake of clarity, we summarize the results per effect. A significant effect of age on mean vowel dispersion was observed (mVCD, $F_{(1–61)}$=4.41, p=.04, R^2=0.07). Contrary to our hypothesis, the speech of older adults was associated with less dispersion for each vowel category (β=−0.26). Neither age, singing practice nor the interaction between these factors influenced any other spectral metric in male speakers. There was, however, an effect of auditory acuity on several metrics (aFCR $F_{(1–61)}$=4.60, p=.034, R^2=0.07; VDI $F_{(1–61)}$=5.76, p=.019, R^2=0.09; mVSD $F_{(1–61)}$=4.20, p=.045, R^2=0.06,; Pillai score $F_{(1–61)}$=11.31, p=.001, R^2=0.16). Surprisingly, higher better ear PTA thresholds were associated with a vowel system expansion (decrease of aFCR scores, β=−0.27), a greater vowel distinctiveness (increase of VDI scores, β=0.29), greater vowel system dispersion (increase in mVSD scores, β=0.25) and less overlap between neighbouring vowels (increase in Pillai score, β=0.39). A significant effect for articulation rate was found on the tongue height metric (F1RR $F_{(1)}$=5.10, p=.027, R^2=0.08), indicating that faster rate is associated with a reduced range of tongue movement (β=-0.28). Effect sizes as measured by Cohen's d indicated strong effects (d>0.8, see Supplementary Material 3).

Turning to female speakers, the F1RR-based model was the best in explaining the data with four significant predictors accounting for 26% of variance: age, singing practice, articulation rate and auditory acuity ($F_{(4–78)}$=8.27, p=.000, adj-R^2=0.26). To compare, the pSVA model with two main effects, age and articulation rate, accounted for 14% of the variance ($F_{(2–80)}$=7.67, p=.001, adj-R^2=14). None of the other models (and respective metrics) yielded significant results. In terms of significant predictors of vowel quality in female speakers, the analyses confirmed an effect of age on two spectral metrics, pVSA and F1RR, and an effect of singing practice on F1 range (F1RR), but no interaction between these factors. Overall, in female speakers, a reduction in vowel articulatory working space was associated with older age (β=-0.40, $F_{(1)}$=3.99, p=.049). An additional analysis showed that the vowel space was considerably larger from young (20–29 years) to middle-aged adults (40–49 years), and smaller following from the fifth decade (Fig. 1, top panel). Similarly, tongue movement along the height dimension was gradually reduced with age (β=-0.23, $F_{(1)}$=6.54, p=.013), whereas tongue movement along the front/back dimension was not significantly related to aging (p=.607). All the above effects were between medium (0.8>d>0.5) and strong (d > 0.8, see Supplementary material 3).

Importantly, the analyses showed that regular singing activity significantly increased tongue movement range along the height dimension in female speakers, independently of their age (β=0.22, $F_{(1)}$=8.99, p=.004, see Fig. 2). Additional analyses showed a significant negative relation between auditory acuity and the range of tongue movement (F1RR β=-0.34, $F_{(1)}$= 5.94, p=.017). Further, faster articulatory rate was negatively related to tongue movement range (F1RR β-0.38, $F_{(1)}$=11.70, p=.001) and the total vowel space area (pVSA β=-0.40, $F_{(1)}$=11.35, p=.001).

In terms of impact, F1 range was best predicted by articulation rate (β=-0.38) followed by auditory acuity (β=-0.33), age (β=-0.23) and

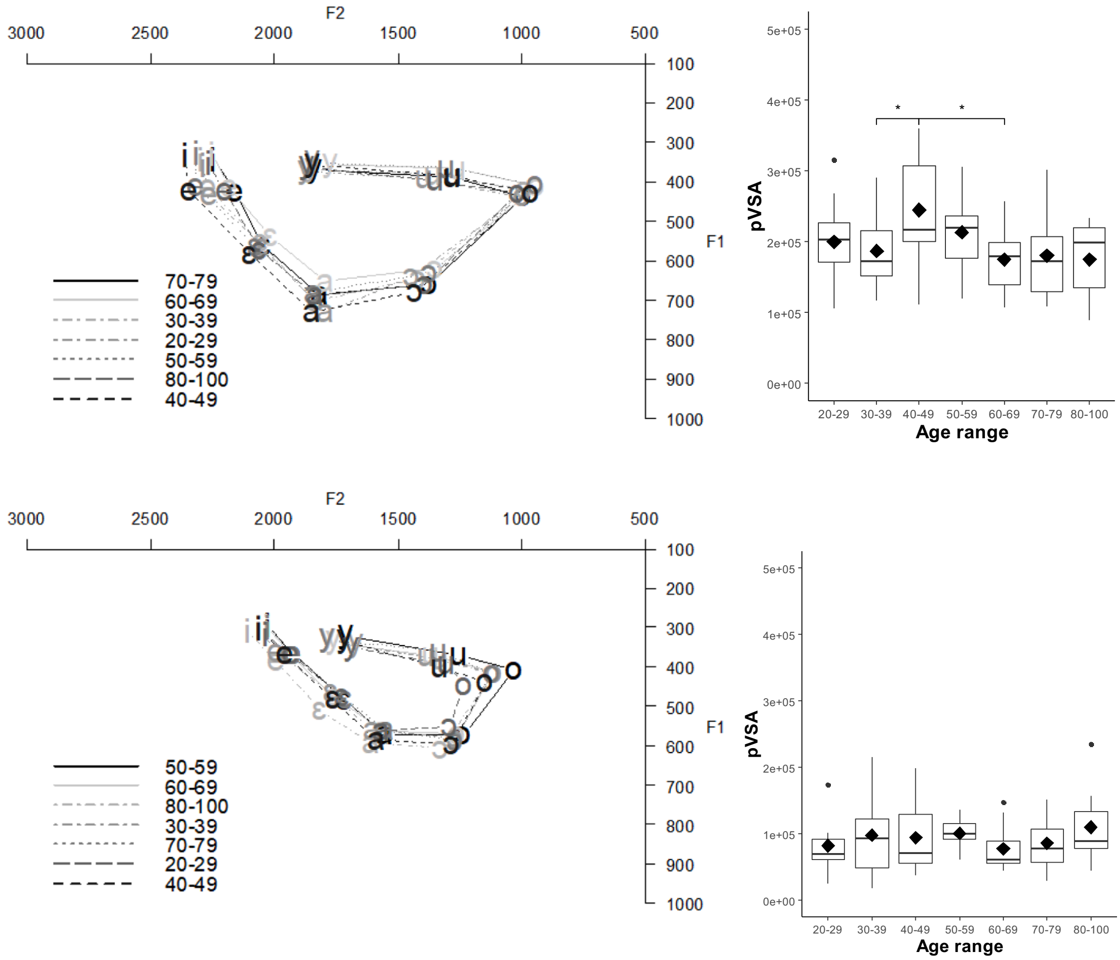


Fig. 1. Variation in age effects on vowel articulatory space in female (top panel) and male speakers (bottom panel) per age range (decades). A tendency toward vowel space shrinking is observed in women. In men, the articulatory working space remains stable across the lifespan (see text).

singing practice (β=0.22). Total vowel area difference was equally well predicted by articulation rate and by the speaker's age (β=-0.40).

3.2. Temporal analyses

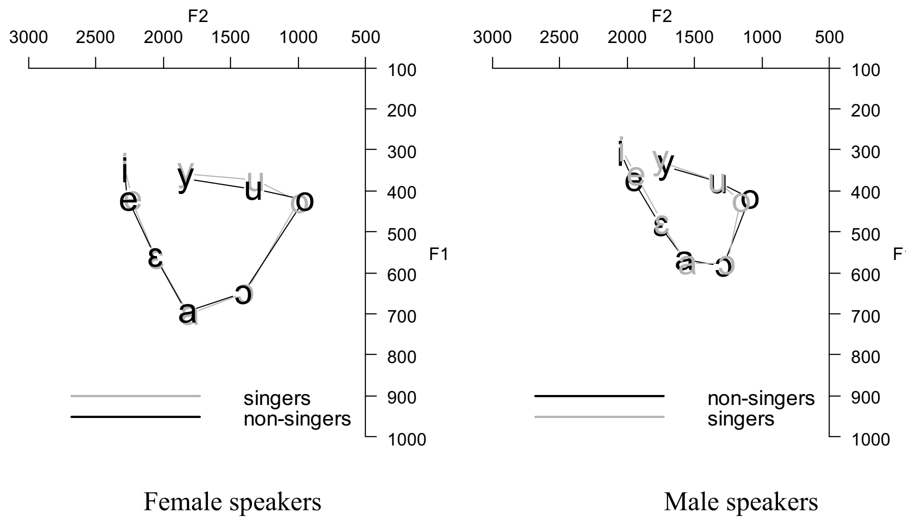
Sex-related differences did not influence temporal metrics (§ 3.1), we thus pooled male and female speakers together for temporal analyses. Regarding the articulation rate, the results of the multiple regression analyses revealed that three variables and an interaction explained 63% of the variance in the data: age, better ear PTA threshold, segment duration and an interaction between age and singing practice ($F_{(4\text{–}141)}$=62.65, p=.000, adj-R^2= 0.63). As expected, articulation rate in read speech was negatively correlated with age (β=-0.36, t=-3.64, p=.000), however, it decreased almost twice as sharply in aging singers as compared to non-singers (β=-0.64, t=-6.91, p = 000 vs. β=-0.36, t = 3.32, p=.001). This is likely related to an inverted pattern of articulatory rates depending on both age (under *vs.* above 50 years) and singing practice, with a singing practice advantage—associated with higher articulation rate—in younger but not in older singers ($F_{(1)}$=4.64, p=.033, see Fig. 3, left panel). Moreover, articulatory rate slightly increased with higher better ear PTA thresholds ($F_{(1)}$=4.46, p=.036, β=0.02) and decreased considerably with longer segment duration ($F_{(1)}$=144.16, p=.000, β=-0.76).

Regarding vocalic segment duration, the results of the multiple linear regression showed that three main effects and an interaction explained 71% of the data ($F_{(5\text{–}140)}$=73.49, p=.000, adj-R^2=0.71). Significant main effects of better ear PTA thresholds ($F_{(1)}$=7.20, p=.008) and age ($F_{(1)}$=175, 14, p=.000) were observed. The analysis of the interaction between age and sex revealed a steeper regression line in male compared to female speakers ($F_{(1)}$=16.47, p=.000, β_{male}=0.64 vs. β_{female}=0.56, see Fig. 4), indicating vowel lengthening across the lifespan being more pronounced in male speakers.

4. Discussion

The primary goal of this study was to examine potential learning effects of singing practice on vowel acoustic characteristics in aging speakers under the general hypothesis that regular singing practice would enhance vowel quality and reduce age-induced changes in speech. The hypothesis was contextualized within opposing accounts on nonspeech-to-speech learning transfer, the *task-dependent* and *task-independent* accounts. To this end, a multiparametric spectral and temporal vowel assessment was carried out on a corpus of 146 adults aged 20 to 98. Taken together, the results indicate that male and female speakers retain overall good vowel targeting in read speech across the lifespan and that singing-to-speech learning effects are very limited.

The age-related acoustic patterns observed are sex specific. Accordingly, female speakers exhibited the expected age effects on vowel quality on two out of six metrics used, the total articulatory working space area (pVSA) and the range of tongue movement along the high-low dimension (F1RR). Both metrics captured a tendency towards vowel space compression with increasing age. A more detailed analysis showed that the maximum vowel area was attained in the fifth decade of life (40–49) and progressively contracted afterwards (Fig. 1, top panel),



Vowel	Non-singers	Singers	Non-singers	Singers
a	699 (3) [690-702]	700 (2.79) [695-706]	567 (2.6) [562-572]	577 (2.99) [571-582]
i	344 (3.70) [337-352]	335 (3.42) [328-342]	303 (3.14) [297-309]	303 (3.14) [297-309]
u	395 (6.77) [382-408]	374 (6.26) [361-386]	379 (5.97) [367-391]	377 (6.68) [364-390]

Fig. 2. Top panel: Vowel space area constructed from mean corner vowels in female (left) and male speakers (right) as a function of singing practice (singers are marked in gray). Bottom panel: mean vowel F1 for three corner vowels /i, a, u/ composing the F1RR metric, on which singing practice effects have been registered for amateur singing female participants.

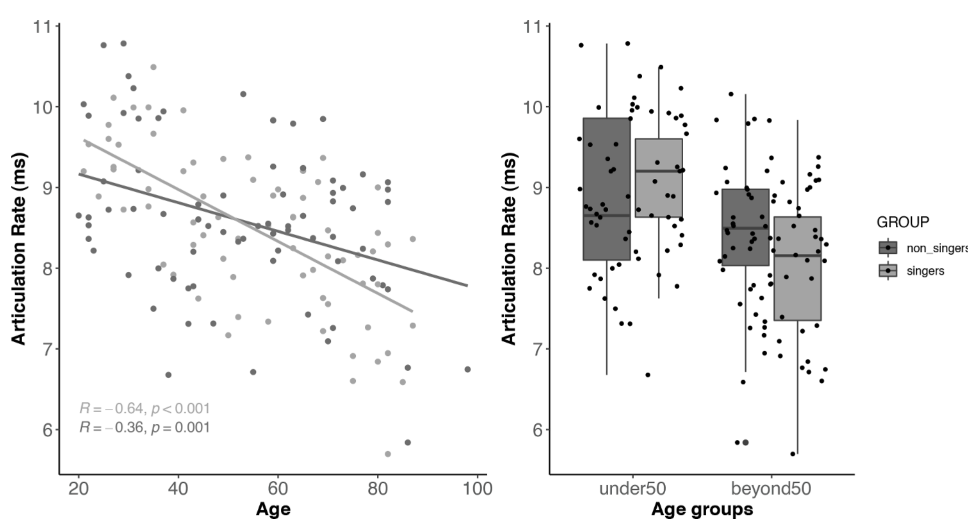


Fig. 3. Negative correlation between articulation rate and age (left) and mean articulation rate for speakers per age group and singing practice (right).

a trend that has been previously observed for women (Gahl and Baayen, 2019). Interestingly, the lack of significant effects on vowel distinctiveness metrics (*e.g.*, VDI, Pillai score) indicates that the reduction of vowel space does not automatically entail a significant contrast loss and suggests a *functional* reorganization within the vowel system that occurs with age. As per male speakers, the overall articulatory vowel space and vowel distinctiveness remained stable across several decades. Contrary to our prediction, vowels produced by older males showed less variability as compared to younger males. Lesser dispersion around the mean can be interpreted as a sign of a mature motor control system and has been previously reported in childhood *vs.* adulthood vowel comparisons (Ménard et al., 2007). An alternative explanation is that older males develop different strategies compared to older females to maintain communication efficiency with declining resources, consistent with the Selection-Optimization-Compensation (SOC) model of aging, which suggests that older adults adjust their objectives and develop compensation strategies to optimize outcomes (Baltes and Carstensen, 1996; Baltes and Lindenberger, 1997; Baltes et al., 1999). The specific mechanics of these processes remains, however, to be investigated.

Turning to singing-to-speech learning transfer, female speakers showed limited benefits of their singing practice, which took the form of increasing the range of tongue movement along the height dimension,

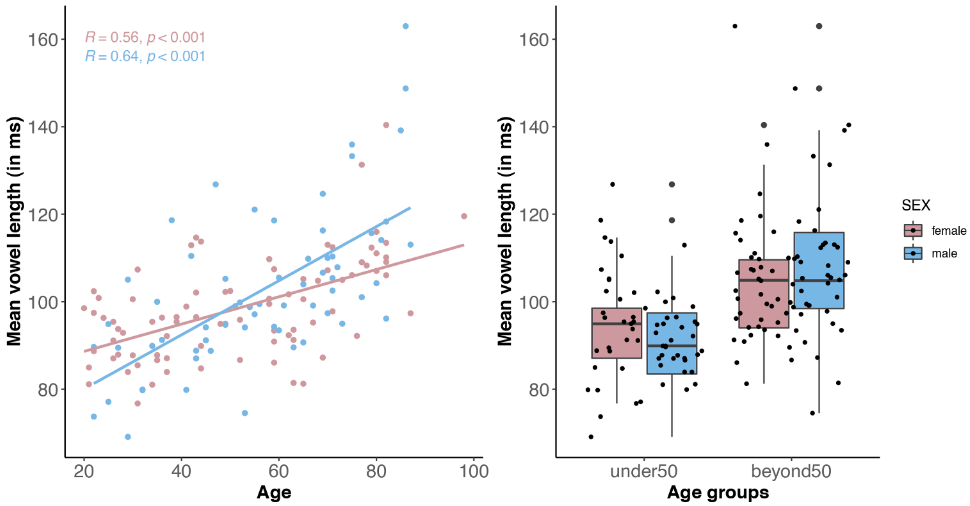


Fig. 4. Correlation between mean vowel duration and age in male and female speakers (left) and mean articulation vowel duration per age group in male and female speakers (right).

which—as shown in the current study—tends to reduce with age (β=-0.23, p=.013). The fact that singing is associated with a reversal of an age-related difference (vertical shrinking of the vowel space) suggests an articulatory gain. Let us recall that F1RR is a ratio-based metric that compares F1 values of three corner vowels /i, a, u/. Optimal score implies lower F1 values for /i/ and /u/ and higher F1 values for /a/. As can be seen in Fig. 2, an increase in the height feature range meant in practice that vowels produced with an elevated tongue /i/ and /u/ were more prototypical (more "closed" *i.e.*, produced with a lower first formant) while the low vowel /a/ remained largely stable. This gain, however, had no impact on vowel distinctiveness, as revealed by the analyses of distinctiveness indexes (*e.g.*, Pillai score). This result reflects the absence of a *linguistic* benefit in singers, as predicted by the task-dependent account. However, this result does not invalidate the task-independent account. One important reason is that there was no motivation for learning, with no significant decrease in vowel quality as a function of age. Another is that change might have been captured on other parameters, for example, dynamic metrics of formant transitions, not included in the study. The result could also be related to the type of material used. Indeed, the main limitation of this study is related to the speech material used. Although read speech is considered more natural than other elicitation tasks such as sustained vowel repetition or isolated word reading, it is not articulatory challenging and thus, may not trigger acoustic change, compensatory effects nor enhance learning. Spontaneous speech may bring into play articulatory rate effects on vowel precision or age effect on vowel distinctiveness, such as those depicted in Fig. 5, and thus would provide complementary information about the aging of articulation in singers and non-singers. Regarding male speakers, none of the spectral metrics registered differences that were attributable to singing practice, which suggests that singing practice did not improve vowel quality in male speakers, at least in read speech.

In terms of temporal control of speech, the analyses showed that aging had similar impact on temporal variables in men and women, with slower articulation and increased vowel duration in older speakers.

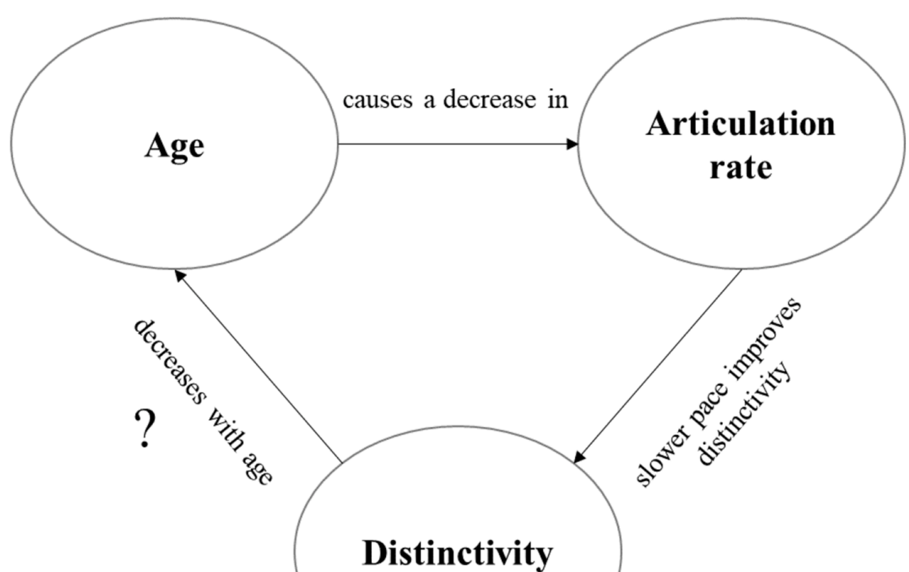


Fig. 5. Three-way circular relationship between articulation rate, vowel distinctiveness and age.

Observed sex differences were related to the magnitude of these effects, not to their direction. Singing practice was associated with greater variability in articulation rates in both groups of speakers. While young to middle-aged adult singers spoke with a faster pace than non-singers of the same age, a reverse trend was observed for older speakers. If age-related segment lengthening and speech tempo decrease are unsurprising and consistent with previous studies (Jacewicz et al., 2009; Linville, 1996), the absence of a significant relationship between temporal factors and distinctiveness metrics is striking. Generally speaking, the interaction between articulatory rate, vowel distinctiveness and age can be defined by a circular three-way dependency (see schematic depiction in Fig. 5). Speech tempo diminishes with age, so does speech intelligibility. However, slower articulation—typically associated with age—favours speech precision and distinctiveness, as established by a wealth of research on vowel quality across different speaking styles and speech rates (*e.g.* Ferguson 2002, Meunier 2012, Moon 1994). While slowing down articulation might be an effect of aging, it may also reflect an adaptive or compensatory strategy to prevent intelligibility loss (Fletcher et al. 2015). Yet, this pattern of results did not emerge from our data. While we did find a general tendency in older speakers to speak at a slower pace, speech tempo did not influence spectral metrics (except for F1RR). Similarly, vowel distinctiveness—as revealed by a signal-based, acoustic assessment typically correlated with intelligibility—was not diminished with age.

Finally, the analyses revealed effects of auditory acuity on vowel characteristics. We interpret these findings within a model that conceptualizes sensorimotor interactions in speech production, the DIVA model (Guenther, 1994). According to DIVA, speakers' auditory acuity should be negatively correlated with vowel category dispersion, since speakers with high acuity have narrower, more focal auditory goals to which they fine-tune their vocalic production (Guenther et al., 2006; Perkell, 2012). Any vowel realization that falls outside of this target region is thus quickly intercepted by the monitoring system and corrected (Martin et al., 2018), which increases articulatory precision and vowel distinctiveness. The DIVA model also accounts for a no-effect of auditory acuity, *i.e.*, no decrease or increase in vowel dispersion metrics, as the vowel parameters stabilize with age and motor system maturation, when the auditory-articulatory tuning becomes less relevant. In this context, the greater vowel distinctiveness associated with lesser auditory acuity we found in male speakers is inconsistent with the model's predictions. Yet this effect was found on four independent metrics of vowel quality: vowel distinctiveness and overlap (VDI, Pillai score) and vowel centralization metrics (aFCR and mVSD). On a closer reading, however, we observe that (i) vowel *category* dispersion (mVCD) is *not* increased in male speakers with less hearing acuity, and that (ii) greater vowel system leads to less overlap between vocalic categories, and thus, better vowel distinctiveness. It thus seems that a lesser reliance on precise auditory feedback—due to a decrease in auditory acuity—has consequences not on isolated acoustic goals for vowels, but rather on the system as a whole. As such, better vowel distinctiveness would be a by-product of the vowel space expansion. This interpretation is supported by the finding that no such effects were observed in female speakers, who exhibited vowel spaces twice as large. Indeed, in the latter group, auditory acuity was significantly associated with shrinking of the vowel space (F1RR), a process typically associated with aging.

In sum, our data support previous findings, specifically, that (i) age effects on vowel quality are predominantly observed in female speakers (Eichhorn et al., 2018; Gahl and Baayen, 2019), and (ii) there is no linear, invariant aging vowel pattern. Further, our data partially support the age-related centralization phenomena (in women). They do not, however, support the hypotheses of age-induced vocal tract lengthening (Endres et al., 1971; Reubold et al., 2010) nor vowel distinctiveness loss.

The novel contribution of this study is two-fold. First, our findings support the idea that articulatory working space does not necessarily determine vowel distinctiveness. This point deserves a more in-depth investigation. The present research is limited to the analysis of the acoustic output, a correlational study between acoustic and perceptual evaluations would shed new light on the relationship between articulatory-acoustic factors in vowel production and their perceptual assessment.

The second original finding is related to the amateur singing practice advantage on expanding the vowel space along the tongue height feature (*i.e.*, F1). Specifically, it was found that female singers tended to lower their first formants as compared with non-singers. Interestingly, this occurred in closed (that is, with an inherently low F1) and lip-rounded vowels (/y/ and /u/). From a musical perspective, such acoustic shifting would allow timbral modulations by tuning the first formant to a lowered fundamental frequency (as in mezzo-soprano). From a linguistic perspective, this effect is inconsistent with an earlier perceptual study (Decoster and Debruyne, 2000) reporting that (i) F2 (*i.e.*, tongue front/back movement) and not F1 was a more robust predictor of age perception, and that (ii) a decrease in F1 (towards the periphery) and not an increase (towards the system's center) was associated with the perception of an 'older' speakers. Our evidence on F1 effects in female singers raises a very interesting question about a possible mediation of the fundamental frequency (F0) in this effect. A study of aging vowels in British English (Reubold et al., 2010) has interpreted a systematic decrease of F1 in aging speakers as a compensatory adjustment to maintain a constant perceptual distance between F0 (known to decrease in elderly speakers, *e.g.*, Lortie et al. (2015) and F1. The fact that we observed this effect in female but not in male speakers -who tend to have stable F0 values across the lifespan (*e.g.* Goy et al., 2013; Stathoppoulos et al., 2011)- speaks to the plausibility of such an interpretation of our finding. Finally, female singers may apply a well-learned vocal tract adjustment used during singing, which consists in avoiding phonation at frequencies around the F1 at 300 Hz, that characterizes closed vowels, to maintain high intensity (Story, 2004). An examination of F0-F1 coupling in female speakers as a function of the singing experience would allow to understand the mechanisms behind this effect. Additionally, an analysis of consonant articulation, as well as both vowel and consonant production in spontaneous speech would provide complementary information that is needed to have a more comprehensive view of the normal aging of speech production mechanisms. Given that age effects on vowel distinctiveness were limited, so were the effects of singing. It is possible that amateur singing improves other aspects of speech and voice functions in aging, but additional investigations are needed.

CRediT authorship contribution statement

Anna Marczyk: Conceptualization, Methodology, Formal analysis, Visualization, Writing – original draft. **Émilie Belley:** Investigation, Formal analysis, Writing – review & editing. **Catherine Savard:** Formal analysis, Writing – review & editing. **Johanna-Pascale Roy:** Conceptualization, Writing – review & editing. **Josée Vaillancourt:** Conceptualization, Writing – review & editing. **Pascale Tremblay:** Conceptualization, Funding acquisition, Methodology, Investigation, Supervision, Resources, Project administration, Writing – review & editing, Data curation.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

We thank two anonymous reviewers for their comments and suggestions that improved the current paper. This work was supported by grants from the Social Sciences and Humanities Research Council of Canada [CRSH; #435-2017-1050 to P.T.], the Drummond Foundation

[2016RFA proposal #27 to P.T.], and the Canadian Foundation for Innovation [31408 to P.T.]; and two Globalink research internships from MITACS. PT holds a Career award from the Fonds de la Recherche en Santé du Québec [FRQ-S, #35016]. We thank Julie Poulin, Lisa-Marie Deschênes, Anne-Christine Bricaud, Elena Vaccaro, Maxime Perron and Antoine Halbaut for their contributions to the recruitment and testing of participants, as well as all others research assistants and research interns who contributed to the project. We also thank Benjamin Côté for his help with vowel data analysis.

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.specom.2022.05.001.

References

Albert, M.L., Sparks, R.W., Helm, N.A., 1973. Melodic intonation therapy for aphasia. Arch. Neurol. 29, 130–131. https://doi.org/10.1001/archneur.1973.00490260074018.
Albouy, P., Benjamin, L., Morillon, B., Zatorre, R.J., 2020. Distinct sensitivity to spectrotemporal modulation supports brain asymmetry for speech and melody. Science 367 (6481), 1043–1047. https://doi.org/10.1126/science.aaz3468.
Apted, M., 1977. *21 Up,* ITV Granada television. Documentary series.
Apted, M. (1984). *28 Up,* ITV Granada television. Documentary series.
Apted, M. (1991). *35 Up,* ITV Gradana television. Documentary series.
Apted, M. (1998). *42 Up,* BBC. Documentary series.
Audibert, N., Fougeron, C., 2012. Distorsions de l'espace vocalique : quelles mesures? Application à la dysarthrie. Actes des XXIXes Journées d'Etudes sur la parole. JEP, Grenoble, France.
Ballard, K.J., Robin, D.A., Folkins, J.W., 2003. An integrative model of speech motor control: a response to Ziegler. Aphasiology 17 (1), 37–48.
Baltes, M.M., Carstensen, L.L., 1996. The process of successful ageing. Ageing Soc. 16, 397–422.
Baltes, P.B., Lindenberger, U., 1997. Emergence of a powerful connection between sensory and cognitive functions across the adult life span: a new window to the study of cognitive aging? Psychol. Aging 12 (1), 12–21. http://www.ncbi.nlm.nih.gov/pubmed/9100264.
Baltes, P.B., Staudinger, U.M., Lindenberger, U., 1999. Lifespan psychology: theory and application to intellectual functioning. Annu. Rev. Psychol. 50, 471–507. https://doi.org/10.1146/annurev.psych.50.1.471.
Benguerel, A.P., Ulkrainetz Mcfadden, T., 1989. The effect of coarticulation on the role of transitions in vowel perception. Phonetica 46 (1–3), 80–96.
Besson, M., Chobert, J., Marie, C., 2011. Transfer of training between music and speech: common processing, attention, and memory. Front. Psychol. 2, 94. https://doi.org/10.3389/fpsyg.2011.00094.
Bilodeau-Mercure, M., Kirouac, V., Langlois, N., Ouellet, C., Gasse, I., Tremblay, P., 2015. Movement sequencing in normal aging: speech, oro-facial and finger movements. Age (Omaha) 37 (4), 37–78.
Bilodeau-Mercure, M., Tremblay, P., 2016. Speech production in aging: linguistic and physiological factors. J. Am. Geriatr. Soc. 64 (11), e177–e182. https://doi.org/10.1111/jgs.14491.
Boersma, P., & Weenink, D. (2011). *Praat: doing phonetics by computer.* (Version 5.2.10) Computer software.
Bunton, K., 2008. Speech *versus* nonspeech: different tasks, different neural organization. Semin. Speech Lang. 29 (4), 267–275. https://doi.org/10.1055/s-0028-1103390.
Carhart, R., Jerger, J., 1959. Preferred method for clinical determination of pure-tone thresholds. J. Speech Hear. Disord. 24 (4), 330–345.
Christiner, M., Reiterer, S.M., 2013. Song and speech: examining the link between singing talent and speech imitation ability. Front. Psychol. 4, 874. https://doi.org/10.3389/fpsyg.2013.00874.
Chung, H., Kong, E.J., Edwards, J., Weismer, G., Fourakis, M., Hwang, Y., 2012. Cross-linguistic studies of children's and adults' vowel spaces. J. Acoust. Soc. Am. 131 (1), 442–454. https://doi.org/10.1121/1.3651823.
Collister, L.B., Huron, D., 2008. Comparison of word intelligibility in spoken and sung phrases. Empir. Musicol. Rev. 3 (3), 109–125.
Decoster, W., Debruyne, F., 2000. Longitudinal voice changes: facts and interpretation. J. Voice 14 (2), 184–193.
Deme, A., 2014. Intelligibility of sung vowels: the effect of consonantal context and the onset of voicing. J. Voice 28 (4), 523. https://doi.org/10.1016/j.jvoice.2014.01.003 e519-525.
Dick, A.S., Bernal, B., Tremblay, P., 2014. The language connectome: new pathways, new concepts. Neuroscientist 20 (5), 453–467. https://doi.org/10.1177/1073858413513502.
Dick, A.S., Tremblay, P., 2012. Beyond the arcuate fasciculus: consensus and controversy in the connectional anatomy of language. Brain 135 (Pt 12), 3529–3550. https://doi.org/10.1093/brain/aws222.
Eichhorn, J.T., Kent, R.D., Austin, D., Vorperian, H.K., 2018. Effects of aging on vocal fundamental frequency and vowel formants in men and women. J. Voice 32 (5). https://doi.org/10.1016/j.jvoice.2017.08.003, 644.e641-644.e649.
Endres, W., Bambach, W., Flosser, G., 1971. Voice spectrograms as afunction of age, voice disguise, and voice imitation. J. Acoust. Soc. Am. 49, 1842–1847.
Ferguson, S.H., Kewley-Port, D., 2002. Vowel intelligibility in clear and conversational speech for normal-hearing and hearing-impaired listeners. J. Acoust. Soc. Am. 112, 259–271.
Fletcher, A.R., McAuliffe, M.J., Lansford, K.L., Liss, J.M., 2015. The relationship between speech segment duration and vowel centralization in a group of older speakers. J. Acoust. Soc. Am. 138 (4), 2132–2139.
Fox, R.A., Jacewicz, E., 2017. Reconceptualizing the vowel space in analyzing regional dialect variation and sound change in American English. J. Acoust. Soc. Am. 142 (1), 444. https://doi.org/10.1121/1.4991021.
Gahl, S., Cibelli, E., Hall, K., Sprouse, R., 2014. The "Up" corpus: a corpus of speech samples across adulthood. Corpus Linguist. Linguist. Theory 10 (2), 315–328. https://doi.org/10.1515/cllt-2013-0023.
Gahl, S.G., Baayen, R.H., 2019. Twenty-eight years of vowels: tracking phonetic variation through young to middle age adulthood. J. Phon. 74, 42–54.
Gendrot, C., Adda.-Decker, M, 2005. Impact of duration on F1/F2 formant values of oral vowels: an automatic analysis of large broadcast news corpora in French and German. In: INTERSPEECH 2005 - Eurospeech, 9th European Conference on Speech Communication and Technology. Lisbon, Portugal.
Gollan, T.H., Goldrick, M., 2018. Aging deficits in naturalistic speech production and monitoring revealed through reading aloud. Psychol. Aging. https://doi.org/10.1037/pag0000296.
Goy, H., Fernandes, D.N., Pichora-Fuller, M.K., van Lieshout, P., 2013. Normative voice data for younger and older adults. J. Voice 27 (5), 545–555. https://doi.org/10.1016/j.jvoice.2013.03.002.
Gregg, J.W., Scherer, R.C., 2006. Vowel intelligibility in classical singing. J. Voice 20 (2), 198–210. https://doi.org/10.1016/j.jvoice.2005.01.007.
Guenther, F.H., 1994. A neural network model of speech acquisition and motor equivalent speech production. Biol. Cybern. 72 (1), 43–53. http://www.ncbi.nlm.nih.gov/pubmed/7880914.
Guenther, F.H., Ghosh, S.S., Tourville, J.A., 2006. Neural modeling and imaging of the cortical interactions underlying syllable production. Brain Lang. 96 (3), 280–301. https://doi.org/10.1016/j.bandl.2005.06.001. S0093-934X(05)00115-X [pii].
Halwani, G.F., Loui, P., Ruber, T., Schlaug, G., 2011. Effects of practice and experience on the arcuate fasciculus: comparing singers, instrumentalists, and non-musicians. Front. Psychol. 2, 156. https://doi.org/10.3389/fpsyg.2011.00156.
Hay, J., Warren, P., Drager, K., 2006. Factors influencing speech perception in the context of a merger-in-progress. J. Phon. 34, 458–484.
Hebert, S., Peretz, I., 2001. Are text and tune of familiar songs separable by brain damage? Brain Cogn. 46 (1–2), 169–175. https://doi.org/10.1016/s0278-2626(01)80058-0.
Hollien, H., Mendes-Schwartz, A.P., Nielsen, K., 2000. Perceptual confusions of high-pitched sung vowels. J. Voice 14 (2), 287–298. https://doi.org/10.1016/s0892-1997(00)80038-7.
Hong, X., Mitchell, R.J., 2007. Backward elimination model construction for regression and classification using leave-one-out criteria. Int. J. Syst. Sci. 38, 101–113. https://doi.org/10.1080/00207720601051463.
Huet, K., Harmegnies, B., 2000. Contribution à la quantification du degré d'organisation des systèmesvocaliques. Actes des XXIIIes Journées d'Etudes sur la parole. JEP, Aussois, France.
Jacewicz, E., Fox, R.A., O'Neill, C., Salmons, J., 2009. Articulation rate across dialect, age, and gender. Lang Var Change 21 (2), 233–256. https://doi.org/10.1017/S0954394509990093.
Jacewicz, E., Fox, R.A., Wei, L., 2010. Between-speaker and within-speaker variation in speech tempo of American English. J. Acoust. Soc. Am. 128 (2), 839–850. https://doi.org/10.1121/1.3459842.
Kelley, M., Tucker, B., 2020. A comparison of four vowel overlap measures. J. Acoust. Soc. Am. 147 (1), 137. https://doi.org/10.1121/10.0000494.
Kent, R.D., Kim, Y.J., 2003. Toward an acoustic typology of motor speech disorders. Clin. Linguist. Phon. 17 (6), 427–445. https://www.ncbi.nlm.nih.gov/pubmed/14564830.
Kent, R.D., Weismer, G., Kent, J.F., Rosenbek, J.C., 1989. Toward phonetic intelligibility testing in dysarthria. J. Speech Hear. Disord. 54 (4), 482–499. https://doi.org/10.1044/jshd.5404.482.
Kim, H., Hasegawa-Johnson, M., Perlman, A., 2012. Vowel contrast and speech intelligibility in dysarthria. Folia Phoniatr. Logop. 63 (4), 187–194. https://doi.org/10.1159/000318881.
Kleber, B., Veit, R., Birbaumer, N., Gruzelier, J., Lotze, M., 2010. The brain of opera singers: experience-dependent changes in functional activation. Cereb. Cortex 20 (5), 1144–1152. https://doi.org/10.1093/cercor/bhp177.
Kuruvilla-Dugdale, M., Dietrich, M., McKinley, J.D., Deroche, C., 2020. An exploratory model of speech intelligibility for healthy aging based on phonatory and articulatory measures. J. Commun. Disord. 87, 105995 https://doi.org/10.1016/j.jcomdis.2020.105995.
Laver, J.T.P., 1991. Phonetic and linguistic markers in speech. In: Laver, J. (Ed.), The Gift of Speech: Readings in the Analysis of Speech and Voice. Edinburgh University Press, pp. 235–264.
Lenth, R., Singmann, H., Love, J., Buerkner, P., Herve, M., 2020. *emmeans: estimated marginal means, aka least-squares means (Version 1.4.5)*. Computer software.
Linville, S.E., 1996. The sound of senescence. J. Voice 10 (2), 190–200.
Lortie, C.L., Thibeault, M., Guitton, M.J., Tremblay, P., 2015. Effects of age on the amplitude, frequency and perceived quality of voice. Age (Dordr) 37 (6), 117. https://doi.org/10.1007/s11357-015-9854-1.

Luo, J., Kothari, N.B., Moss, C.F., 2017. Sensorimotor integration on a rapid time scale. Proc. Natl. Acad. Sci. USA 114 (25), 6605–6610. https://doi.org/10.1073/pnas.1702671114.
Maas, E., 2017. Speech and nonspeech: what are we talking about? Int J Speech Lang Pathol 19 (4), 345–359. https://doi.org/10.1080/17549507.2016.1221995.
Martel Sauvageau, V., Macoir, J., Langlois, M., Prud'Homme, M., Cantin, L., Roy, J.P., 2014. Changes in vowel articulation with subthalamic nucleus deep brain stimulation in dysarthric speakers with Parkinson's disease. Parkinsons Dis 2014, 487035. https://doi.org/10.1155/2014/487035.
Martin, C.D., Niziolek, C.A., Dunabeitia, J.A., Perez, A., Hernandez, D., Carreiras, M., Houde, J.F., 2018. Online adaptation to altered auditory feedback is predicted by auditory acuity and not by domain-general executive control resources. Front Hum Neurosci 12, 91. https://doi.org/10.3389/fnhum.2018.00091.
McCloy, D.R. (2016). *phonR: tools for phoneticians and phonologists.* (Version 1.0-7). Computer software.
Ménard, L., Schwartz, J.L., Boë, L.J., Aubin, J., 2007. Articulatory-acoustic relationships during vocal tract growth for French vowels: analysis of real data and simulations with an articulatory model. J. Phon. 35, 1–19.
Merrett, D.L., Peretz, I., Wilson, S.J., 2013. Moderating variables of music training-induced neuroplasticity: a review and discussion. Front. Psychol. 4, 606. https://doi.org/10.3389/fpsyg.2013.00606.
Meunier, C., Espesser, R., 2012. Vowel reduction in conversational speech in French: the role of lexical factors. J. Phon. 39 (3), 271–278.
Meunier, C., Ghio, A., 2018. Caractériser la distinctivité du système vocalique des locuteurs. Acte XXXIIe Journées d'Études sur la Parole. JEP, Aix-en-Provence, France.
Moon, S.J., Lindblom, B., 1994. Interaction between duration, context, and speaking style in English stressed vowels. J. Acoust. Soc. Am. 96, 40–55. https://doi.org/10.1121/1.410492.
Morris, R., Brown, W.S., 1987. Age-related voice measures among adult women. J. Voice 1 (1), 43.
Nasreddine, Z.S., Chertkow, H., Phillips, N., Bergman, H., Whitehead, V., 2003. Sensitivity and specificity of the montreal cognitive assessment (MoCA) for detection of mild cognitive deficits. Can. J. Neurol. Sci. 30, 30.
Nycz, J., Hall-Lew, L., 2014. Best practices in measuring vowel merger. In: *Proceedings of the Meetings on Acoustics: The Acoustical Society of America through the American Institute of Physics* (1 ed., Vol. 20, pp. 1-19). https://doi.org/10.1121/1.4894063.
Ozdemir, E., Norton, A., Schlaug, G., 2006. Shared and distinct neural correlates of singing and speaking. Neuroimage 33 (2), 628–635. https://doi.org/10.1016/j.neuroimage.2006.07.013.
Padovani, M., Gielow, I., Behlau, M., 2009. Phonarticulatory diadochokinesis in young and elderly individuals. Arq. Neuropsiquiatr. 67 (1), 58–61. https://www.ncbi.nlm.nih.gov/pubmed/19330213.
Patel, A.D., 2011. Why would musical training benefit the neural encoding of speech? The OPERA hypothesis. Front. Psychol. 2, 142. https://doi.org/10.3389/fpsyg.2011.00142.
Patel, A.D., 2012. The OPERA hypothesis: assumptions and clarifications. Ann. N. Y. Acad. Sci. 1252, 124–128. https://doi.org/10.1111/j.1749-6632.2011.06426.x.
Patel, A.D., 2014. Can nonlinguistic musical training change the way the brain processes speech? The expanded OPERA hypothesis. Hear. Res. 308, 98–108. https://doi.org/10.1016/j.heares.2013.08.011.
Patel, A.D., Peretz, I., 1997. Is music autonomous from language? A neuropsychological appraisal. In: Deliège, I., Sloboda, J. (Eds.), Perception and Cognition of Music. Erlbaum, pp. 191–215.
Perkell, J., Guenther, F., Lane, H., Matthies, M.L., Perrier, P., Vick, J., Wilhelms-Tricarico, R., Zandipour, M., 2000. A theory of speech motor control and supporting data from speakers with normal hearing and with profound hearing loss. J. Phon. 28, 233–272. https://doi.org/10.1006/jpho.2000.0116.
Perkell, J.S., 2012. Movement goals and feedback and feedforward control mechanisms in speech production. J. Neurolinguistics 25 (5), 382–407. https://doi.org/10.1016/j.jneuroling.2010.02.011.
Perkell, J.S., Lane, H., Denny, M., Matthies, M.L., Tiede, M., Zandipour, M., Vick, J., Burton, E., 2007. Time course of speech changes in response to unanticipated short-term changes in hearing state. J. Acoust. Soc. Am. 121 (4), 2296–2311. http://www.ncbi.nlm.nih.gov/pubmed/17471743.
Perron, M., Theaud, G., Descoteaux, M., Tremblay, P., 2021. The fronto-temporal organization of the arcuate fasciculus and its relationship with speech perception in young and older amateur singers and non-singers. Hum. Brain Mapp. 42 (10), 3058–3076. https://doi.org/10.1002/hbm.25416 https://doi.org/.
Perron, M., Vaillancourt, J., Tremblay, P., 2022. Amateur singing benefits speech perception in aging under certain conditions of practice: behavioural and neurobiological mechanisms. Brain Struct. Funct. 227 (3), 943–962. https://doi.org/10.1007/s00429-021-02433-2 https://doi.org/.
R Core Team, 2019. *R: a language and environment for statistical computing.* R Foundation for Statistical Computing, Vienna, Austria.
Ravi, S.K., Shabnam, S., George, K.S., Saraswathi, T., 2019. Acoustic and aerodynamic characteristics of choral singers. J. Voice 33, 5. https://doi.org/10.1016/j.jvoice.2018.03.018, 803 e801-803 e805.
Reubold, U., Harrington, J., Kleber, F., 2010. Vocal aging effects on F0 and the first formant: a longitudinal analysis in adult speakers. Speech Commun. 52 (7–8), 638–651. https://doi.org/10.1016/j.specom.2010.02.012.
Rothman, K.J., 1990. No adjustments are needed for multiple comparisons. Epidemiology 1 (1), 43–46.
Sadagopan, N., Smith, A., 2013. Age differences in speech motor performance on a novel speech task. J. Speech Lang. Hear. Res. 56 (5), 1552–1566. https://doi.org/10.1044/1092-4388(2013/12-0293.
Sapir, S., Ramig, L.O., Spielman, J.L., Fox, C., 2010. Formant centralization ratio: a proposal for a new acoustic measure of dysarthric speech. J. Speech Lang. Hear. Res. 53 (1), 114–125. https://doi.org/10.1044/1092-4388(2009/08-0184.
Saville, D.J., 1990. Multiple comparison procedures: the practical solution. Am. Stat. 44, 174–180. https://doi.org/10.1080/00031305.1990.10475712.
Skodda, S., Grönheit, W., Schlegel, U., 2012. Impairment of vowel articulation as a possible marker of disease progression in Parkinson's disease. PLoS One 7 (2), e32132. https://doi.org/10.1371/journal.pone.0032132.
Smith, B.L., Wasowicz, J., Preston, J., 1987. Temporal characteristics of the speech of normal elderly adults [Research support, U.S. Gov't, P.H.S.]. J. Speech. Hear. Res. 30 (4), 522–529. http://www.ncbi.nlm.nih.gov/pubmed/3695445.
Souza, P.E., Arehart, K.H., Shen, J., Anderson, M., Kates, J.M., 2015. Working memory and intelligibility of hearing-aid processed speech. Front. Psychol. 6, 526. https://doi.org/10.3389/fpsyg.2015.00526.
Sparks, R.W., Helm, N.A., Albert, M.L., 1974. Aphasia rehabilitation resulting from melodic intonation therapy. Cortex 10, 303–316. https://doi.org/10.1016/S0010-9452(74)80024-9.
Stathoppoulos, E., Huber, J.E., Sussman, J.E., 2011. Changes in acoustic characteristics of the voice across the life span: measures from individuals 4–93 years of age. J. Speech Lang. Hear. Res. 54 (4), 1011–1021. https://doi.org/10.1044/1092-4388(2010/10-0036.
Story, B.H., 2004. Vowel acoustics for speaking and singing. Acta Acust. United Acust. 90 (4), 629–640.
Tobias, J.V., 1977. The need for preventive audiology. Audiol. Hear. Educ. 3 (2), 37–38.
Tremblay, P., Deschamps, I., 2016. Structural brain aging and speech production: a surface-based brain morphometry study. Brain Struct. Funct. 221 (6), 3275–3299. https://doi.org/10.1007/s00429-015-1100-1.
Tremblay, P., Deschamps, I., Bedard, P., Tessier, M.H., Carrier, M., Thibeault, M., 2018. Aging of speech production, from articulatory accuracy to motor timing. Psychol. 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. Hum. Brain Mapp. 38 (5), 2751–2771. https://doi.org/10.1002/hbm.23558.
Tremblay, P., Veilleux, J.-A., 2018. The impact of singing on human communication in aging: from protection to rehabilitation. In: Heydon, R., Fancourt, D., Cohen, A. (Eds.), Routledge Companion to Interdisciplinary Studies in Singing. Volume III Well-Being. Taylor & Francis Books.
Van Der Merwe, A., 2021. New perspectives on speech motor planning and programming in the context of the four- level model and its implications for understanding the pathophysiology underlying apraxia of speech and other motor speech disorders. Aphasiology 35 (4), 397–423. https://doi.org/10.1080/02687038.2020.1765306.
Vorperian, H.K., Kent, R.D., 2007. Vowel acoustic space development in children: a synthesis of acoustic and anatomic data. J. Speech Lang. Hear. Res. 50 (6), 1510–1545. https://doi.org/10.1044/1092-4388(2007/104.
Weismer, G., 2006. Philosophy of research in motor speech disorders. Clin. Linguist. Phonet. 20, 315–349.
Wickham, H., 2016. ggplot2: Elegant Graphics For Data Analysis. Springer-Verlag. http://ggplot2.org.
Yesavage, J.A., Brink, T.L., Rose, T.L., Lum, O., Huang, V., Adey, M., Leirer, V.O., 1982. Development and validation of a geriatric depression screening scale: a preliminary report [Research support, U.S. Gov't, Non-P.H.S.]. J. Psychiatr. Res. 17 (1), 37–49. http://www.ncbi.nlm.nih.gov/pubmed/7183759.
Ziegler, W., 2002. Task-related factors in oral motor control: speech and oral diadochokinesis in dysarthria and apraxia of speech. Brain Lang. 80 (3), 556–575. https://doi.org/10.1006/brln.2001.2614.
Zumbansen, A., Tremblay, P., 2018. Music-based interventions for aphasia could act through a motor-speech mechanism: a systematic review and case–control analysis of published individual participant data. Aphasiology 33 (4), 466–497.