

JOURNAL ARTICLE

# Using naïve listener imitations of native speaker productions to investigate mechanisms of listener-based sound change

Christopher Carignan<sup>1,2</sup>

<sup>1</sup> Institut für Phonetik und Sprachverarbeitung, Ludwig-Maximilians-Universität München, Munich, DE

<sup>2</sup> The MARCS Institute for Brain, Behaviour and Development, Western Sydney University, Sydney, AU  
[c.carignan@phonetik.uni-muenchen.de](mailto:c.carignan@phonetik.uni-muenchen.de)

This study was designed to test whether listener-based sound change—listener misperception (Ohala, 1981, 1993) and perceptual cue re-weighting (Beddor, 2009, 2012)—can be observed synchronically in a laboratory setting. Co-registered articulatory data (degree of nasalization, tongue height, breathiness) and acoustic data (F1 frequency) related to the productions of phonemic oral and nasal vowels of Southern French were first collected from four native speakers, and the acoustic recordings were subsequently presented to nine Australian English naïve listeners, who were instructed to imitate the native productions. During these imitations, similar articulatory and acoustic data were collected in order to compare the articulatory strategies used by the two groups. The results suggest that the imitators successfully reproduced the acoustic distinctions made by the native speakers, but that they did so using different articulatory strategies. The articulatory strategies for the vowel pair /ã/-/a/ suggest that listeners (at least partially) misperceived F1-lowering due to nasalization and breathiness as being due to tongue height. Additional evidence supports perceptual cue re-weighting, in that the naïve imitators employed nasalance less, and tongue height more, in order to obtain the same F1 nasal-oral distinctions that the native speakers had originally produced.

**Keywords:** sound change; imitation; perception; acoustics; articulation; nasalization; breathiness; ultrasound; EGG; nasalance; co-registration

## 1. Introduction

### 1.1. Listener-based sound change processes

The phonetic influence of sounds that are adjacent to one another in the temporal speech chain is generally referred to as co-articulation. Sound change processes that arise due to these influences are typologically common and are easy to explain on the level of production. For example, the fronting of a velar tongue closure before a front vowel—e.g., Latin *civitātem* [ki:wita:tem] ‘city’ > Italian *città* [tʃit:a], French *cit * [site]—is a common sound change pattern among the world’s languages, and it is readily explained on an articulatory basis: The context of the palatal constriction for [i] gives rise to an articulatory constraint on the velar constriction for [k], resulting in palatalization of the velar consonant. The change is considered to be dependent upon the conditioning environment and it is generally assumed to originate from the speaker herself—after all, it is the speaker’s own vocal tract that provides the context for the articulatory constraint (cf. Guion, 1998).

However, some sound changes are not articulatorily motivated in this way. Take, for example, the change from labio-velar stops in Proto-Indo-European to velar stops in

Classical Greek, as shown in 1.1 (Meillet, 1967), and from velar stops in Proto-Bantu to labial stops in West Teke, as shown in 1.2 (Guthrie, 1967–1970).<sup>1</sup>

1.1	Proto-Indo-European	Classical Greek	
	* <i>ekwos</i>	<i>hippos</i>	‘horse’
	* <i>gwiwos</i>	<i>bios</i>	‘life’
1.2	Proto-Bantu	West Teke	
	* <i>-kumu</i>	<i>pfumu</i>	‘chief’

Changes such as these are not considered to originate from the speaker’s own articulatory constraints, since the lips and the tongue dorsum are under independent articulatory control and are not expected to influence one another. Another example involves changes between nasalization and glottalization (laryngealization), and between glottalization and nasalization, e.g., certain dialects of British English [hã:f] ‘half’ (Matisoff, 1970). The link between these articulatorily unconnected features (termed *rhinoglottophilia* by Matisoff, 1975) is not considered to arise from co-articulatory influence, since articulatory control of velum opening/closing is independent of articulatory control of vocal fold abduction/adduction. Similarly, cases of ‘spontaneous nasalization’ have been observed diachronically, wherein phonologically distinctive nasalization has emerged in environments that were once adjacent to high-airflow, spread-glottis consonants (Bloch, 1920, 1965; Grierson, 1922; Ohala, 1975; Ohala & Ohala, 1993; Turner, 1921)—e.g., Sanskrit *sarpa* ‘snake’ > Hindi [sãp]. Ohala (1975) argued a perceptual basis for these changes: When the glottis is spread, the lowered amplitude and increased bandwidth of F1 created by the coupling of the oral and subglottal cavities can ‘mock’ comparable acoustic effects due to coupling of the oral and nasal cavities, and listeners can misperceive and lexically re-analyze these effects accordingly. Ohala and Busa (1995) later argued that these same perceptual factors can explain the diachronic ‘inverse’ of spontaneous nasalization: sound changes involving the loss of nasals before voiceless fricatives.

In a similar manner to the link between nasalization and glottalization, sound change patterns involving the link between nasalization and tongue height modification have been observed. Diachronically, systematic changes in vowel height have been documented typologically as nasal vowel systems develop: High nasal vowels have a tendency to be produced with a lower tongue position and low nasal vowels have a tendency to be produced with a higher tongue position as vowel nasality becomes phonologized (Beddor, 1982; Hajek, 1997; Sampson, 1999). A notable example comes from the evolution of nasal vowels in French, e.g., the lowering of [i] in Latin *vinum* [winum] ‘wine’ > Old French *vin* [vĩn] > Modern French *vin* [vẽ] (Sampson, 1999). Although there does exist an intrinsic muscular connection between the soft palate and the tongue via the *palatoglossus* muscle (Kuehn & Azzam, 1978; Zemlin, 1998), this connection is argued to affect velic height when the tongue lowers, but not tongue height when the velum lowers. When considering a possible effect of velum lowering on tongue height, the two articulations are generally considered to be under independent control.

## 1.2. Listener-based sound change mechanisms

Since sound changes like these do not have any apparent articulatory motivation, they are considered to arise from the listener rather than the speaker. But what are possible mechanisms by which listener-based sound change could occur? Two prominent theories of listener-based sound change mechanisms are discussed in the current section: listeners’

<sup>1</sup> Both examples are reproduced from Ohala (1989, p. 182), 4 (a–b).

misperception of the source of acoustic effects (Ohala, 1981, 1993) and listeners' re-weighting of co-varying articulatory properties (Beddor, 2009, 2012).

According to Ohala (1981, 1993), listener-based sound change can be accounted for in terms of listener misperception of the articulatory source of acoustic modulations. Under normal circumstances, listeners are able to correctly attribute the effects of temporally overlapping gestures to their respective articulatory sources (i.e., perceptual compensation; Mann & Repp, 1980; Yu & Lee, 2014; Zellou, 2017). However, in some cases, the listener may fail to compensate for this temporal overlap and, thus, fail to attribute various effects to their proper sources. In these cases, the listener misperceives the effects that arise from one articulatory source as being due to a different articulatory source. Thus, according to Ohala, the change itself first happens at the level of speech perception, when the listener has failed to properly contextualize the various properties of the speech signal (Harrington, Kleber, Reubold, Schiel, & Stevens, in press). When the listener becomes the speaker, this re-analysis of the articulatory source is passed to the level of production, and the listener-turned-speaker subsequently produces the novel articulatory variant in her own speech.

Beddor (2009, 2012) posits a slightly different mechanism for listener-based sound change, one which also assumes that the acoustic consequences of co-articulated speech variants provide the necessary 'raw material' (Beddor, 2009, p. 787) for the types of sound change discussed in this paper. The difference in Beddor's model is that listeners are sensitive to the acoustic consequences of temporally overlapping articulatory gestures, and they do not need to misperceive these consequences in order for sound change to occur. In other words, whereas Ohala argues that listeners' access to phonetic details is typically due to misperception of the articulatory source, Beddor argues that listeners regularly access these details in order to facilitate perception. However, a given listener may nevertheless differ from a given speaker with respect to the co-articulatory source and effect that is represented in their respective grammars. According to Beddor, an important first stage of sound change that may result from this process is when an articulatory source and its phonetic effect begin to exhibit a trading relationship with one another. When this happens, a listener may correctly perceive a phonetic effect, but she may attribute it to a segment other than its original source. An example given by Beddor involves phonetically nasalized  $[\tilde{V}N]$  sequences, wherein the listener may correctly perceive the phonetic effect  $[\tilde{V}]$ , but fail to parse all of the effect with the speaker source  $/N/$ , and instead partially contribute the effect to the adjacent segment  $/V/$ . In this way, Beddor's theory of listener-based sound change does not exclude the possibility of listener misperception—i.e., misperceptions may happen, but sound change can occur even when speech is accurately perceived. Instead, listener-based sound change in Beddor's model occurs when listeners perceive a speech utterance with a different weighting between co-varying articulatory properties than was originally produced by the speaker. In other words, co-varying articulations aren't necessarily misperceived, they are simply weighted in a different manner by the listener than by the speaker.

### **1.3. Testing listener-based sound change mechanisms**

The current study has been created to observe possible evidences of these listener-based sound change mechanisms in the laboratory, using naïve listener imitations of native speaker productions. Specifically, Australian English naïve listeners were recruited to imitate phonemic nasal-oral vowel pair tokens produced by Southern French native speakers.

These two particular languages were chosen because vowel nasality is phonologically contrastive in (Southern) French (i.e.,  $/\tilde{V}/$ ), while it is phonetically contextual in

(Australian) English (i.e., [VN]). Thus, when imitating the native French tokens, the English listeners would be unable to perceptually compensate and associate nasality with a specific source, N, because there is no context for contextual normalization to occur (i.e., no final N in such French words). If articulatory re-analysis is indeed involved in listener-based sound change mechanisms in some manner, removing the source of a specific effect will help ensure that such re-analysis might be observed in an experimental setting. Furthermore, when imitating productions of phonemic nasal vowels, English speakers would not be biased by their own phonological representation of vowel nasality in their native language, as opposed to, e.g., Hindi or Portuguese speakers.

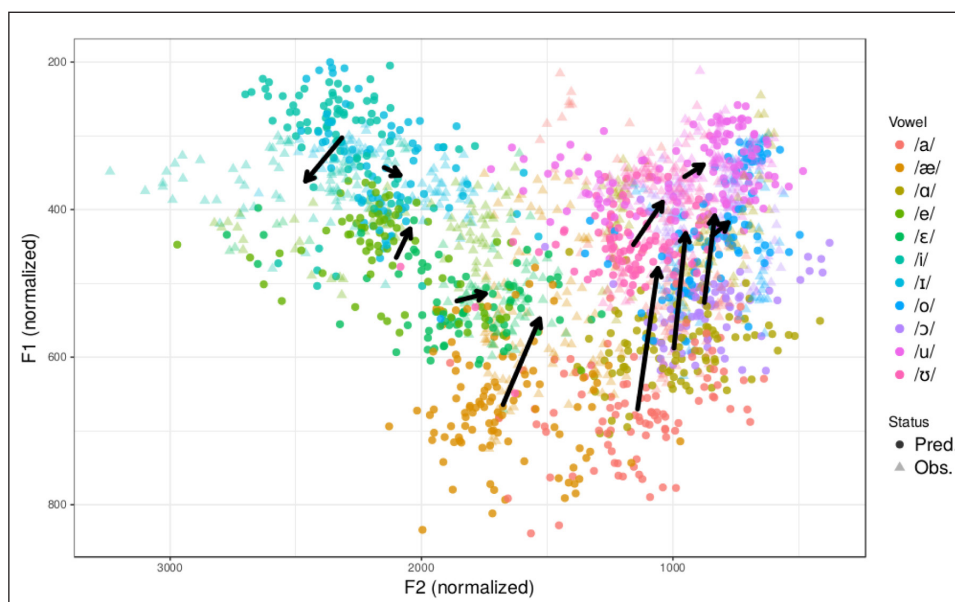
Naïve listeners were recruited for the imitation task instead of native speakers (e.g., speakers from a different variety of French) or L2 learners of French, since both of these speaker groups would be influenced by their knowledge of the language. Furthermore, a specific prediction from Ohala (1993) is that less experienced listeners are likely to be the primary drivers of sound change, given that they are presumed to have a lesser ability to perceptually compensate for the articulatory source of acoustic effects. Included in this ‘less experienced’ group are listeners such as children or language learners; I contend that naïve listeners are even less experienced than these, and that the use of naïve listeners in an imitation task will further help to ensure that the participants fail to perceptually compensate in this way.

Vowel nasalization, specifically, has been chosen for this study due to the inherently ambiguous nature associated with its production. In the following section, ambiguities associated with the overlapping effects of nasalization, tongue height change, and breathy voicing on F1 frequency are detailed. These ambiguities make vowel nasalization an ideal test case for investigating listener-based sound change: The possibility of multiple many-to-one articulatory-to-acoustic mappings creates the possibility of multiple sources of articulatory re-analysis by listeners.

### 1.3.1. The multi-dimensional nature of vowel nasality

In the production of vowel nasality, the velum is lowered in order to couple the oral and nasal cavities via the velo-pharyngeal port—known as *VP coupling* or, simply, nasalization. Due to the coupling of the two cavities, nasalization results in a wide range of acoustic effects in comparison to non-VP-coupled oral vowels. While these include reduced acoustic energy and increased formant bandwidths in low frequencies (Stevens, 2000, p. 193), variation in F1 frequency is also observed due to the introduction of additional pole-zero pairs to the acoustic spectrum (Maeda, 1993). The direction of the shift of F1 frequency depends on the respective frequencies of the oral poles, nasal poles, and nasal zeros, as well as the degree of VP coupling itself. Accordingly, it is not always practical to attempt to determine F1 of the oral cavity alone for VP-coupled vowels. Therefore, the designation F1' will be used in this manuscript to refer to the most prominent low-frequency spectral peak of a VP-coupled vowel, regardless of whether it arises from the oral tract, the nasal tract, or the combination of the two. Recent articulatory and acoustic evidence has shown that the most pervasive F1'-frequency-related consequence of VP coupling is lowering of F1' for non-front, non-high vowels (Carignan, 2018). **Figure 1** displays the direction of shift in F1/F2 frequencies due to the independent effects of VP coupling, after correction for effects due to oral configuration; the arrows display the average direction of the independent effects on F1/F2 frequencies for each vowel category.

However, VP coupling is not the only possible source of F1' frequency modulation in VP-coupled vowels. Perhaps the most familiar source of F1 modulation for vowels (be they oral or nasal) is variation in tongue height. Generally, a higher tongue body is correlated with a lower F1 frequency, while a lower tongue body is correlated with a



**Figure 1:** Independent effects of nasalization on acoustic vowel quality. Arrows indicate the direction of the effects in F1/F2 space. Replicated from Carignan (2018).

higher F1 frequency (Johnson, 2003; Stevens, 2000). Thus, differences in tongue height are predicted to modify F1' frequency of VP-coupled vowels in a manner that is independent of VP coupling itself. A spread glottis can also have an independent effect on the spectral characteristics of F1', due to the coupling of the oral cavity to the subglottal cavity, as previously mentioned. For example, breathy voice is differentiated from modal voice by a more prominent first harmonic (H1) and a lower-amplitude, wider-bandwidth F1 (Bickley, 1982; Garellek, 2014; Garellek & Keating, 2011; Gordon & Ladefoged, 2001; Hanson, K. N. Stevens, Kuo, Chen, & Slifka, 2001; Kreiman et al., 2012; Lotto, Holt, & Kluender, 1997). Since H1 has, by definition, a low frequency, the combination of a higher-amplitude H1 and a lower-amplitude F1 is generally predicted to lower F1 center of gravity, which can subsequently affect F1' frequency. In particular, the increased energy from H1 is expected to decrease F1' frequency, especially for non-high vowels, for which the oral F1 is higher in frequency than H1. Likewise, VP coupling has the effect of increasing the prominence of the first harmonic, decreasing the amplitude of F1, and increasing the bandwidth of F1 for nasal vowels compared to oral vowels (Chen, 1997; Maeda, 1993; Styler, 2015, 2017), as well as lowering F1' frequency for non-high vowels (Carignan, 2018).

Thus, in the production of vowel nasalization, there exists the possibility for VP coupling, tongue height, and breathiness to affect F1' frequency in ways that are independent but nevertheless similar to one another. In this manner, F1' frequency modulation is unique among the acoustic characteristics of nasalization, which is why focus has been placed on F1' in the current study, rather than on other acoustic characteristics of nasalization (e.g., changes in tongue height are not predicted to affect F1 bandwidth). Vowel nasalization thus creates the potential for interaction between these articulatory variables in the perceptual domain, i.e., due to misperception of the articulatory source of the F1' modulation: If F1' is lowered for a given nasal vowel, can listeners ascertain whether the F1'-lowering is due to tongue height change, nasalization, or breathiness? Evidence of this perceptual interaction is described in the following section.

### 1.3.2. Co-variation of nasalization, tongue height, breathiness, and F1' in perception

Kingston and Macmillan (1995) observed that F1' and nasality were integrated in listeners' perception of synthesized vowels: The perception of nasality was influenced by F1', and the

perception of F1' was influenced by nasality. Beddor, Krakow, and Goldstein (1986) and Krakow, Beddor, and Goldstein (1988) observed that the F1' variation due to nasalization can be attributed by listeners to changes in tongue height for non-contextually nasalized vowels. Of particular interest to the current study is their finding that a decrease in F1' for low vowels can be attributed to either a higher tongue position or an increase in degree of nasalization. Similarly, Wright (1975, 1986) found that listeners perceived nasalized [ã] as higher than oral [a]. These findings suggest the F1'-lowering for non-high vowels that is due to the independent effects of velum lowering (Carignan, 2018) can, in fact, be perceived by listeners as a raising of the tongue.

In Ohala and Amador (1981), perceptual stimuli were created by iterating acoustic periods adjacent to voiceless fricatives and acoustic periods adjacent to nasal consonants, using audio recordings of vowel productions from American English and Mexican Spanish. These stimuli were perceived by listeners as having the same degree of nasalization, even though co-articulatory nasalization was only present in one of the two environments. Similar results were replicated for Hindi using the same methodology (Ohala & Ohala, 1992). These results suggest that the acoustic effects of a spread glottis (e.g., voiceless fricatives, breathy voicing) can be perceived as nasalization. Using audio from speakers with cleft palate, Imatomi (2005) created stimuli with six degrees of breathiness and filtered with three degrees of simulated hypernasality (mild, moderate, severe). She found that the more breathy stimuli increased listeners' judgment of hypernasality for mild hypernasal filters, but decreased judgment of hypernasality for moderate and severe hypernasal filters. Finally, Lotto et al. (1997) observed that listeners perceived synthesized breathy low vowels as higher than non-breathy low vowels in English.

These results suggest that the perception of nasalization, breathiness, tongue height, and F1' exist in a naturally co-varying relationship: (1) breathy voice can be perceived as nasalization, (2) breathy voicing of low vowels can be perceived as tongue raising, (3) nasalization of non-high (including low) vowels can be perceived as tongue raising, and (4) F1'-lowering for nasalized low vowels can be perceived as either tongue raising or increased nasalization.

### 1.3.3. Co-variation of nasalization, tongue height, breathiness, and F1' in production

As the term *nasal* implies, the primary articulatory distinction between a nasal vowel and an oral vowel is the relative height of the velum. However, in addition to velum lowering, phonemic nasal vowels have been observed to be produced with modifications to tongue posture in Northern French (Bothorel, Simon, Wioland, & Zerling, 1986; Brichler-Labaeye, 1970; Carignan, 2014; Carignan, Shosted, Fu, Z.-P., & Sutton, 2015; Delvaux, 2012; Delvaux, Metens, & Soquet, 2002; Demolin, Delvaux, Metens, & Soquet, 2003; Straka, 1965; Zerling, 1984), Southern French (Carignan, 2017), Laurentian French (Carignan, 2013), Brazilian Portuguese (Barlaz et al., 2015; da Matta Machado, 1993; Shosted, 2015; Shosted et al., 2015), European Portuguese (Martins, Oliveira, Silva, & Teixeira, 2012; Oliveira, Martins, Silva, & Teixeira, 2012; Teixeira et al., 2012), and Hindi (Shosted, Carignan, & Rong, 2012). Similar lingual differences have been observed for phonetically nasalized vowels, as well. In American English, for example, evidence from electromagnetic articulography (EMA) has shown that speakers raise the tongue body during the production of phonetically nasalized oral /i/ (Carignan, Shosted, Shih, & Rong, 2011) and/or lower the tongue body during the production of phonetically nasalized oral /a/ (Arai, 2004). The authors of both studies posited that these modifications to tongue height may represent articulatory compensation for changes in F1' frequency that arise

from VP coupling and, therefore, may be used by speakers in order to help prevent a phonemic nasal-oral split.

Vowel nasality has also been observed to co-vary with changes in voice quality. Larynx lowering has been observed during the production of nasals (Riordan, 1980; Westbury, 1979); lowering the larynx not only has the effect of lowering formant frequencies (in particular, F1 frequency, as noted by Denning, 1989), but also biases phonation towards breathiness due to the lateral force that placed on the vocal folds (Moisik, Lin, & Esling, 2014). Using both electroglottographic and acoustic measurements, Garellek, Ritchart, and Kuang (2016) investigated the co-production of phonetic nasalization and breathiness of vowels in Bo, Luchan Hani, and Southern Yi. They found that contextually nasalized vowels were produced with greater breathiness than oral vowels in all three languages. The authors posited that this co-variation has arisen due to listener misperception and/or phonetic enhancement: Due to common acoustic effects for both nasalization and breathiness (discussed previously), it is likely that (1) the acoustic effects of co-articulatory nasalization on the vowel were misperceived as effects of breathiness and, consequently, speakers began producing the vowels with breathiness; and/or (2) speakers produce these phonetically nasalized vowels with breathy voice as a way of enhancing the acoustic characteristics of nasalization.

These results suggest that the production of nasalization, breathiness, tongue height, and F1' exist in a naturally co-varying relationship: (1) vowel nasality can be produced with changes in tongue height, (2) vowel nasality can be produced with breathy voicing, (3) nasalization has a tendency to be produced with a lowered larynx, biasing phonation towards breathiness, and (4) tongue raising, nasalization, and breathiness lower F1' frequency for non-high vowels in independent ways.

#### **1.4. Hypotheses**

In summary, it has been shown in previous research that systematic relationships between nasalization, breathiness, tongue height, and F1' exist at the levels of both production and perception. As such, the current study has been designed to examine the co-variation of these three articulatory variables and their relation to F1' frequency in the productions of the phonemic nasal-oral vowel distinction in Southern French, and in the imitation of these productions by Australian English naïve imitators. Comparison between the native productions and the naïve imitations will serve as a way to test for possible evidence to support the theoretical frameworks for listener-based sound change described in Section 1.2.

Since the relationship between these three articulatory variables and F1' represents a classic *many-to-one* problem of articulation-to-acoustics (i.e., there are multiple articulations that can affect the singular acoustic characteristic), evidence of a *one-to-many* relationship is not expected. In other words, no plausible hypotheses can be made concerning an outcome in which the Australian English (AE) imitators produce the same articulatory distinctions as the Southern French (SF) speakers, yet the F1' distinctions are different for the two groups. Such an outcome would suggest that one (or both) of the groups modify F1' using articulatory distinctions that are not monitored in this study (e.g., lip configuration or pharyngeal aperture), and it is therefore outside of the scope of the particular research questions pursued here. However, a number of hypotheses can be made for a *many-to-one* relationship resulting from the AE imitators producing the same F1' distinctions as the SF speakers but not the same articulatory distinctions. These hypotheses are related to the two theoretical frameworks as follows:

1. Evidence in support of listener misperception (Ohala):
  - a. The SF speakers produced articulatory distinctions that the AE imitators fail to replicate.
  - b. The AE imitators produce articulatory distinctions that the SF speakers had not originally produced.
2. Evidence in support of perceptual cue re-weighting (Beddor):  
The AE imitators produce the same articulatory distinctions as the SF speakers but to different degrees.

Alternatively, it may be the case that the AE imitators produce the same F1' distinctions as the SF speakers, as well as all of the same articulatory distinctions. This outcome would suggest that no mechanism of listener-based change is represented in the listener-turned-speaker imitations and, thus, would fail to provide evidence in support of either of the two theoretical frameworks.

## 2. Methods

The current study was designed to investigate the articulatory co-variation used to produce phonemic vowel nasality, and to compare the articulatory strategies employed by native speakers with those used by naïve imitators. As such, data collection proceeded in two separate stages: an initial stage for native speakers, and a subsequent stage for naïve imitators. In the first stage of the study, articulatory and acoustic data were collected from four male native speakers of SF producing oral (/a,ɛ,o/) and nasal (/ã,ẽ,õ/) vowel counterparts in French lexical items in a carrier phrase. This stage has been described in detail in Carignan (2017), but portions of the study will be summarized here. In the second stage of the study, the same data were collected from nine male AE speakers who had no linguistic experience with a language containing phonemic nasal vowels.<sup>2</sup> For data collection in this second stage, the participants listened through headphones to a randomized selection of word stimuli created from the SF native productions; they were instructed simply to imitate the productions. In the following sections, the speaker groups, word list, stimuli, and data acquisition protocol for the two parts of the study are explained in detail.

### 2.1. Speakers, word list, and stimuli

Only males were included for both the SF native speaker and AE naïve imitator groups in order to control for oral articulatory differences between male and female speakers with regard to nasal vowel articulation (Engwall, Delvaux, & Metens, 2006). For the SF native speaker data collection, the word list contained 16 monosyllabic French CV lexical items. The word list was balanced for C voicing (/p,b/) and target V (/ẽ,ɛ,ã,a,õ,o/). Words with high vowels (/i,u/) were also included as filler items and for normalizing tongue height for each speaker (see Section 3.1.3). The complete word list is available in Table A of Appendix A. 10 blocks of the word list were included, totaling 160 items for each speaker (120 experimental items and 40 filler items). Target words appeared on a computer screen in the carrier phrase *Ils écrivent X parfois* “They write X sometimes.” Bilabial consonants preceding and following the target vowel helped to avoid lingual co-articulatory effects

<sup>2</sup> Specifically, during recruitment, potential participants were asked whether they had ever had any formal training in, or any linguistic experience with French, Portuguese, or Hindi. This question was asked in order to determine if they had had any significant linguistic exposure to phonemic vowel nasality—these three languages being the most likely candidates for such exposure—without drawing the participants’ attention to the specific research aims of the study.



from neighboring consonants. The speakers were instructed to read each carrier phrase at a normal speed and volume, and to advance the slides at their own pace.

The audio file for each SF speaker (see Section 2.2 for recording details) was imported into Praat (Boersma & Weenink, 2015), and the amplitude was normalized by scaling the peak of the entire audio file to 0.99. Normalizing the audio in this way helped to mitigate amplitude variation in the imitation stimuli (e.g., due to variation in speaker amplitude and/or recording levels), while maintaining natural amplitude variation among the target words (e.g., due to differences in oral impedance across vowel qualities). The target words were extracted from the amplitude-normalized audio, using the broadband spectrogram and respective waveform to manually identify the C onset and V offset. The nearest zero-amplitude-crossing point in the waveform was identified for both the onset and offset, and the audio data between these points were extracted and saved as separate .WAV files.

To create the stimulus order for the imitation task, a random selection of three tokens (i.e., from three different blocks) was made for each word and for each SF speaker, so that no duplicate items appeared in a given stimulus set. This resulted in 192 stimuli per set (16 items  $\times$  three tokens  $\times$  four speakers). This randomization was carried out three times, resulting in three separately randomized but equally balanced stimulus sets. The .WAV files corresponding to these stimuli were converted to 192 kbps .MP3 files and embedded into LaTeX .PDF presentation files using the Beamer document class. The presentation files were constructed to play the individual audio files automatically as the participant advanced each slide. Audio was presented to the participants using a Roland Duo-Capture EX USB audio interface and Sennheiser HD-650 headphones. The participants were told that they would hear words in a foreign language, one word per slide, and to simply imitate each word as well as they could. Before each experimental session began, the participant practiced the task on the first few items in order to familiarize himself with the task, and so that the experimenter could adjust the audio to a clear and comfortable level for the participant.

## 2.2. Data acquisition

Data from both groups were collected at the Analysis of Human Articulatory Actions Laboratory of The MARCS Institute for Brain, Behaviour and Development (Western Sydney University). For both SF speakers and AE imitators, two separate computers were used to simultaneously record ultrasound, nasalance, EGG, and acoustic data. A third computer was used to either display the presentation slides (for the native speakers) or to play the embedded audio files (for the imitators). An example of the naïve imitator experimental setup is shown in **Figure 2**.

### 2.2.1. Computer 1: Ultrasound + nasalance

A GE LOGIQ *e* portable ultrasound system was used to image the midsagittal tongue surface. A GE 8C-RS transducer was held in place using a non-metallic elastic headset (Derrick, Best, & Fiasson, 2015), which helped to maintain contact of the probe against the skin while also maximizing speaker comfort and jaw maneuverability. For all speakers, imaging was performed at a frequency of 8 MHz and at a depth between 8–10 cm. In order to align the probe imaging plane with the speaker's midsagittal plane, the speaker was first instructed to produce a range of speech sounds while the experimenter adjusted the position of the transducer underneath the speaker's jaw until there were no longer evidences of plane misalignment (e.g., discontinuity in the tongue surface between the dorsum and the blade). Real-time ultrasound video was captured from the GE LOGIQ *e* VGA video output using an Epiphan VGA2USB Pro video grabber.



**Figure 2:** Experimental equipment setup (shown: naïve imitator data collection re-enactment).

A Glottal Enterprises H-SEP-MU nasalance plate was used to collect nasalance data. When the plate is placed between the speaker's upper lip and nostrils, two microphones located on either side of the plate produce a stereo audio signal where one channel contains audio from the microphone on the upper side of the plate (near the nose) and the other channel contains audio from the microphone on the lower side of the plate (near the mouth), with acoustic separation between the two channels (due to the plate itself acting as an acoustic baffle). Unlike nasal airflow systems, nasalance systems do not create a seal around the mouth or nose, which both acts as a lowpass filter and effectively lengthens the vocal tract (lowering formant frequencies). Thus, nasalance can be measured in an open system like the one displayed in **Figure 2** without significantly altering the acoustic signal, allowing a separate microphone to be used to record high-fidelity audio (see Section 2.2.2). The nasalance audio and ultrasound video data were co-registered on a single computer, using Fast Forward Moving Pictures Expert Group software (FFmpeg; FFmpeg Development Team, 2016) to record an .AVI file with video (30 fps) and embedded audio (44.1 kHz).

### 2.2.2. Computer 2: EGG + audio

EGG is a non-invasive method of indirectly measuring the changing surface contact between the vocal folds during phonation. Two electrodes are secured externally on either side of the larynx, and a low-amplitude current passes from one electrode, through the vocal folds, to the other electrode. As the vocal folds are adducted and abducted, the electrical resistance to this current decreases and increases, respectively, due to the changing area of contact between the vocal folds. A Glottal Enterprises EG2-PCx two-channel electroglottograph was used to record both EGG data and high-fidelity audio. Before each experimental session, the EGG electrodes were cleaned and a conductive gel was applied to the electrode surface. The optimal electrode height was determined by instructing the speaker to sustain the sound /a/ while the electrodes were adjusted vertically along the larynx, using the dedicated larynx height signal on the EG2-PCx to guide placement. Once the optimal height was achieved, the electrodes were secured in place with a Velcro strap. High-fidelity audio was collected with a Behringer ECM 8000 omnidirectional condenser microphone, using the EG2-PCx as a preamplifier. Audacity software (Audacity Development Team, 2016) was used to record both data streams at a sampling rate of 44.1 kHz from the EG2-PCx USB output.

### 2.3. Data synchronization and segmentation

In order to synchronize the four data signals collected on the two separate computers, the signals were segmented manually according to temporally comparable acoustic and articulatory landmarks. At the beginning and end of each experimental session, the

speaker produced the sounds “TA,TA,TA.” Since the sequence [ta] provides both transient acoustic information (a burst) and transient articulatory information (a rapid movement of the tongue blade), it is relatively straightforward to identify corresponding acoustic and articulatory landmarks for this sequence. The ultrasound, nasalance, EGG, and audio data between the two landmarks were extracted for each participant’s recording and, hence, synchronized.

The target vowel intervals were segmented manually in Praat, using the onset of high-amplitude periodicity in the acoustic signal as the onset and the offset of high-amplitude periodicity in the EGG signal as the offset. Using the EGG signal to determine the vowel offset provided a more accurate segmentation of laryngeal phonation compared to using the acoustic signal alone; however, the acoustic signal was nonetheless necessary for determining the vowel onset, since the boundary between [b] and the target V for voiced onset items was indistinguishable in the EGG signal.

### 3. Analyses

#### 3.1. *Articulatory and acoustic measurements*

F1', nasalance, tongue height, and EGG measurements were made at time points relating to 25%, 50%, and 75% of the vowel interval for each token. Since the ultrasound data were recorded at the lowest sampling rate (i.e., 30 Hz compared to 44.1 kHz for all other data), the measurements for all signals were carried out at time points associated with the ultrasound frames: The ultrasound frames nearest 25%, 50%, and 75% of each target vowel interval were identified automatically, and the time points related to these three frames were used for the measurements of all four signals. Some tokens were produced with such a short duration that three ultrasound frames could not be identified within the vowel interval. These tokens, all of which contained oral vowels, were excluded from analysis. This occurred for two native speakers and four naïve imitators, but only for a very small set of tokens in each case: as few as one experimental item for naïve imitator AE01 (0.7% of total) and as many as six experimental items for native speaker SF04 (4.17% of total).

##### 3.1.1. F1'

The high-fidelity audio signals were denoised in Audacity and high-pass filtered at 50 Hz in Praat. F1' values in the filtered signals were then measured automatically using the Burg LPC<sup>3</sup> method in Praat, with an estimation of the first two formants. Although only F1' is included in this study, two-formant estimation was found to yield the most reliable results for F1'. Different maximum formant values were used for each speaker and for each vowel pair, in order to minimize formant estimation error. For the native speaker data, the optimized parameter setting was determined by manually adjusting the maximum formant for each speaker and vowel pair until the automatic formant tracking aligned consistently with the F1' band visible in broadband spectrogram. For the naïve imitator data, the optimized parameter setting was determined using a semi-automated procedure similar to Escudero, Boersma, Rauber, and Bion (2009). Firstly, F1' measurements were taken at the three time points of each token in an iterative manner: At each iteration in the range from 1000 Hz to 3500 Hz, the F2 ceiling was increased by 50

<sup>3</sup> When a pole is introduced to a spectrum near in frequency to another pole, the additional spectral energy can shift the spectral center of gravity of the original pole towards the new pole. Thus, the additional spectral energy from the first nasal pole (N1) in a nasalized vowel, or a more prominent first harmonic (H1) in a breathy vowel, can shift the center of gravity of F1 if N1 and/or H1 are near in frequency to F1. This shift in F1 center of gravity can also influence the measurement of F1, since the overall spectral shape is affected by the presence of N1/H1. Thus, an all-pole formant estimation technique such as LPC may potentially be influenced by N1/H1, even if the oral F1 remains constant.

Hz and formant measurements were logged. Secondly, the set of formant measurements was imported into R (R Core Team, 2016) and the cumulative variance for F1' at the three time points was measured separately for each 50 Hz ceiling step and for each vowel pair. Thirdly, the variance for F1' was summed and plotted for the three time points of each vowel pair; the plots were visually inspected and the ceiling which yielded the lowest variance was logged (i.e., the ceiling parameter which produced the most consistent F1' measurements at all three time points for the given vowel pair). Finally, the suitability of these optimized parameter settings was verified manually by inspecting the resulting formant tracks against a broadband spectrogram in Praat. After each of the optimized F2 ceilings was verified for each vowel pair and for each speaker, the final logged parameters were used to measure F1' at the three time points of each target vowel. The F2 ceilings used in the final analysis are given in Table B of Appendix A.

### 3.1.2 Nasalance

Nasalance is a measure of the proportion of the amplitude (i.e., sound pressure energy) related to air radiated by the nasal tract, relative to the amplitude related to air radiated by both the nasal and oral tracts (i.e.,  $\frac{A_{nasal}}{(A_{oral} + A_{nasal})}$ ). Using the separate acoustic signals from the two microphones on the nasalance separator plate, amplitude tracks for the oral and nasal signals were created using a custom script in Praat, and nasalance values were calculated automatically for the three time points of each token.

### 3.1.3. Tongue height

An estimation of tongue height was made from locating the highest point of contours fitted to the midsagittal tongue shapes in the ultrasound images related to the three time points of each target vowel. Before tongue contour fitting, the ultrasound frames were first rotated to the speaker's occlusal plane. Occlusal plane angles were determined by inserting a wooden tongue depressor between the participant's teeth, at an angle spanning the back molars on one side of the mouth and the incisors on the other side. The speaker was instructed to bite down on the tongue depressor and press his tongue underneath and upwards against it. The area of contact between the tongue and the depressor appeared as a straight line in the ultrasound image, and the angle of this line was used for image rotation in Matlab (The Mathworks Inc., 2015).

Tongue contours were fitted using EdgeTrak software (Li, Kambhamettu, & Stone, 2005). In cases where EdgeTrak was unable to fit contours reliably (usually occurring for the high vowels /i,u/), the tongue contour was fit manually. The fitted contours were exported with 100-point spline interpolation, and the highest y-value for each tongue contour was logged and used as the tongue height metric for the respective ultrasound frame. The high vowel /i,u/ tokens served as maxima in the scaling of tongue height values and, thus, tongue height normalization for each speaker. Scaling was calculated for a range of 0–1, with 0 representing the lowest tongue position produced by a speaker and 1 representing the highest tongue position. Thus, a given value represents a proportion of the entire range of each participant's tongue height values across the vowels /a,ã,ε,ẽ,o,õ,i,u/.

### 3.1.4. Contact quotient

Contact quotient (CQ) refers to the proportion of a glottal cycle in which the vocal folds are in contact with each other. Since breathy voicing is produced with a relatively opened glottis (with respect to modal voicing), it is characterized by a lower proportion of vocal fold contact in a glottal cycle and, thus, lower CQ values (Baken & Orlikoff, 2000; Garellek, 2014). In contrast to acoustic measurements of breathiness—e.g., increased spectral tilt (Bickley, 1982; Garellek, Ritchart, & Kuang, 2016; Garellek, Samlan, Gerratt,

& Kreiman, 2016)—CQ can be used to measure differences in voice quality in a manner that is unaffected by other articulatory changes—e.g., increased spectral tilt resulting from vowel nasalization (Styler, 2015, 2017). Given the nature of the current research (i.e., investigation of independent articulatory variables), independent measurements of each articulator are necessary.

CQ was calculated automatically using a custom Praat script. The EGG signal was first smoothed using a 21-frame sliding triangle window ( $\approx 0.5$  ms). The first derivative of this smoothed signal (DEGG) was calculated, and the DEGG maximum in each glottal cycle was located automatically. Each maximum in the DEGG signal represents the moment the vocal folds start to come together, i.e., the beginning of the closing phase. The end of the closing phase (i.e., the beginning of the opening phase) was calculated according to the Hybrid method (Orlikoff, 1991), defined as the point at which the EGG signal falls below 25% of the amplitude range of the glottal cycle. For each of the three time points in a given target vowel interval, the nearest DEGG maximum was determined and CQ values were calculated for both the glottal cycle preceding and following this maximum. The average of these two CQ values was calculated and used as the measurement for that particular time point.

### **3.2. Nasal-oral token matching**

The measurements outlined above were used to ascertain the overall articulatory and acoustic distinctions between oral and nasal vowel counterparts. However, in order to compare the differences between the SF native speakers and the AE naïve imitators with respect to these distinctions, as well as with respect to how the articulatory distinctions relate to F1' distinction for the two groups, a method similar to bootstrap sampling was used to create randomized matched comparisons between oral and nasal tokens for each participant (i.e., each SF speaker and each AE imitator), according to the following step-wise process. Before creating nasal-oral matched comparisons, the data were normalized by calculating  $z$ -scores for each variable, for each participant. After normalization, for each AE imitator's data set, an oral item was first selected at random. Subsequently, a nasal item was selected at random from among the total set of nasal tokens that matched the selected oral token in terms of vowel pair category, onset consonant, and the SF speaker who produced the token. For example, if the random oral item selected for imitator AE05 was a token of *beau* [bo] “beautiful” produced by speaker SF02, then the random matched nasal item was a token of *bon* [bõ] “good” that was also produced by speaker SF02 and presented during imitator AE05's data collection session. Once the matched selection was made, the speaker-normalized articulatory and acoustic  $z$ -score values for the oral token were subtracted from the same  $z$ -score values for the nasal token for both the AE imitator and the SF speaker. These  $z$ -score differences were logged for both the imitator and the speaker, and the two matched tokens were removed from the data set before the iterative process repeated for a new pair of tokens. This randomized matching process continued until each oral token in the data set was matched with a corresponding nasal token. Since the randomized stimuli sets that were presented to the AE imitators were balanced for both word and SF speaker, this step-wise randomized matching process also resulted in a balanced set of nasal-oral matched tokens for each imitator and speaker.

Using this method, the nasal-oral distinctions for a given variable can be compared between the AE imitators and the SF speakers, in order to observe whether the naïve imitators make the same nasal-oral F1', nasalance, tongue height, and CQ distinctions as the native speakers, as well as whether they make these distinctions to the same degree that the native speakers originally had made. Additionally, each of the articulatory nasal-oral distinctions can be compared to the F1' nasal-oral distinctions, in order to observe

how the articulatory variables contribute to variation in F1', as well as whether these articulatory-acoustic relationships differ between the two groups.

### 3.3. Statistical analyses

Linear mixed-effects (LME) models were created in R using the *lmer* function in the *lme4* package (Bates, Mächler, Bolker, & Walker, 2015). Three sets of models were created: one set to test the F1' and articulatory differences between oral and nasal vowel counterparts (an example of the model structure is provided in 2.1 for the effect of NASALITY on CQ), and two sets to test the nasal-oral differences between the two groups, using the nasal-oral matched measurements described above. Within these latter two sets, one set was created to test the effect of GROUP on the nasal-oral variable distinctions (an example of the model structure is provided in 2.2 for the tongue HEIGHT distinction), and a separate set was created to test the effect of GROUP on the relation between the articulatory variables and F1' (an example of the model structure is provided in 2.3 for the relation between the NASALANCE distinction and the F1' distinction). Separate models for each variable and vowel pair combination were created within each set, resulting in 12 models each for the first and second sets (3 vowel pairs  $\times$  4 variables) and 9 models for the third set (3 vowel pairs  $\times$  3 variables).

2.1  $\text{lmer}(\text{CQ} \sim \text{NASALITY} + (1|\text{IMITATOR}) + (1|\text{SPEAKER}) + (1|\text{ORDER}) + (1|\text{ONSET}))$

2.2  $\text{lmer}(\text{HEIGHT}_{\text{diff}} \sim \text{GROUP} + (1|\text{IMITATOR}) + \dots + (1|\text{ONSET}))$

2.3  $\text{lmer}(\text{F1}_{\text{diff}} \sim \text{NASALANCE}_{\text{diff}} * \text{GROUP} + (1|\text{IMITATOR}) + \dots + (1|\text{ONSET}))$

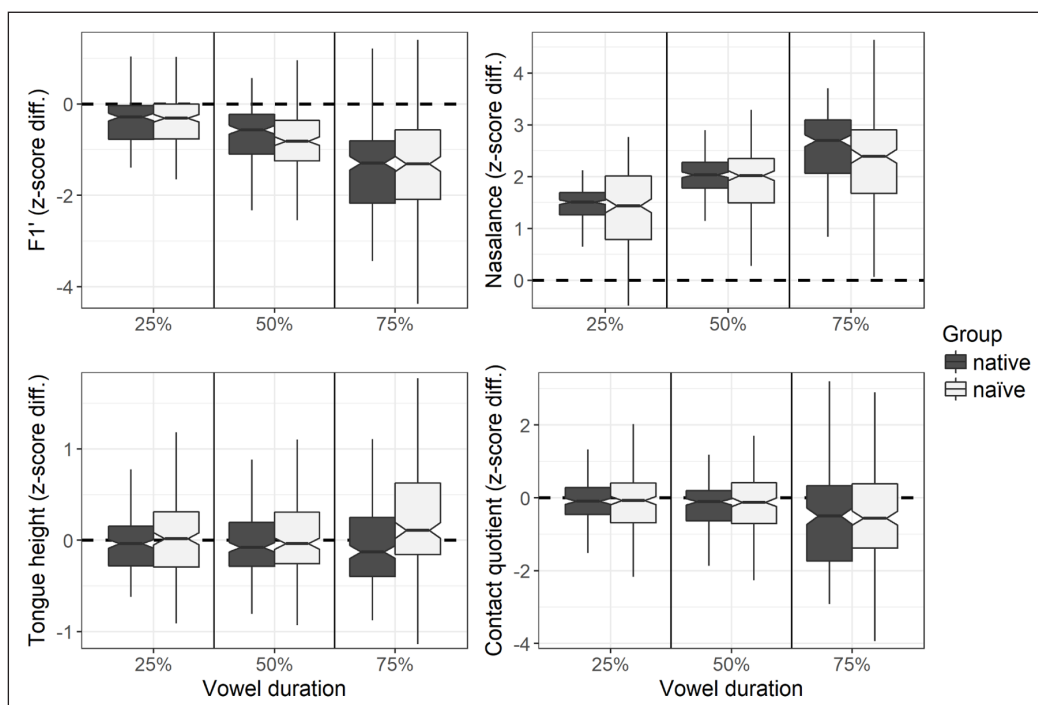
In each model, random intercepts were included for naïve IMITATOR, native SPEAKER, token ORDER, and ONSET consonant. Estimates for degrees of freedom, *t*-statistics, and *p*-values were generated using Satterthwaite approximation under the *lmerTest* package (Kuznetsova, Brockhoff, & Christensen, 2016). For all models, the alpha level was adjusted using Bonferroni correction for comparison across the 12 models in the first and second sets ( $\alpha = 0.004$ ) or across the 9 models in the third set ( $\alpha = 0.006$ ). The complete results are provided in Appendix A: Table C for the first set of models, Table D for the second set, and Table E for the third set.

## 4. Results

### 4.1. Articulatory and F1' nasal-oral distinctions

#### 4.1.1. /ã/ vs. /a/ distinctions

Box plots of /ã/-/a/ z-score differences for F1', nasalance, tongue height, and CQ values are shown in **Figure 3**; individual box plots created for each speaker and imitator are provided in Appendix B. For any given box plot, the horizontal line is the median of the distribution, the center notch pointing inward displays the 95% confidence interval around the median, the box displays the interquartile range (i.e., the middle 50% of the data), and the vertical lines extending from the box demarcate  $1.5 \times$  the interquartile range. Outliers have been removed to aid visualization. With regard to F1', the first LME model reveals that /ã/ was produced with lower F1' values than /a/ throughout the entire vowel (evidenced also by the negative values in the box plots), and the box plots suggest an increase in this frequency difference throughout the vowel interval (i.e., the box plots are positioned at increasingly lower values across the three time points). No significant effect of GROUP was observed in the second LME model, revealing that the speaker-normalized F1' distinction between /ã/ and /a/ was the same for both the naïve imitators and the native speakers. These results suggest that the AE imitators were generally successful in replicating the F1'-frequency-related acoustic pattern produced by the SF speakers.



**Figure 3:** /ã/-/a/ z-score differences for F1', nasalance, tongue height, and CQ, as produced by the SF native speakers (dark box plots) and the AE naïve imitators (light box plots). Separate box plots are displayed for measurements taken at 25%, 50%, and 75% of the vowel duration.

With regard to nasalance, the first LME model reveals that /ã/ was produced with a greater degree of nasalance than /a/ (evidenced also by the positive values in the box plots). The box plots suggest that the nasalance distinction increases in a linear manner throughout the vowel interval for both groups, which is generally characteristic of phonetically co-articulated vowel nasalization (Cohn, 1990). This observation is not surprising for the Australian English imitators, since vowel nasalization is a phonetic phenomenon in English. However, the same pattern observed for the Southern French native speakers is in contrast to the more constant, plateau-like pattern that is characteristic of nasal vowels in Northern Metropolitan French (Cohn, 1990; Delvaux, 2012). The significant (negative) effect of GROUP in the second LME model reveals that the AE imitators produced a smaller nasalance distinction between /ã/ and /a/, in comparison to the SF speakers.

With regard to tongue height, although the first LME model reveals that /ã/ and /a/ were produced with the same overall tongue height, the box plots suggest slightly different patterns for the two groups: Whereas the native speakers produced /ã/ with a slightly lower tongue position than /a/, the naïve imitators produced /ã/ with a slightly higher tongue position than /a/, especially in the latter part of the vowel interval. The second main LME model reveals that this difference between the two groups is indeed significant: The AE naïve imitators produced a higher overall /ã/-/a/ tongue height distinction in comparison to the SF native speakers. Moreover, separate post-hoc LME models with the same random structure as the main models reveal that these separate group-wise patterns are significant for both the SF speakers ( $t(1,1186.5) = -6.35, p < 0.001$ ) and the AE imitators ( $t(1,394.6) = 4.97, p < 0.001$ ).

With regard to CQ, the first LME model reveals that /ã/ was produced with lower CQ values compared to /a/; this is especially apparent in the box plots in the latter part of the vowel interval, suggesting that /ã/ becomes even breathier towards the end of the vowel. No significant effect of GROUP was observed in the second LME model, revealing that the speaker-normalized CQ distinction between /ã/ and /a/ was the same for both the naïve

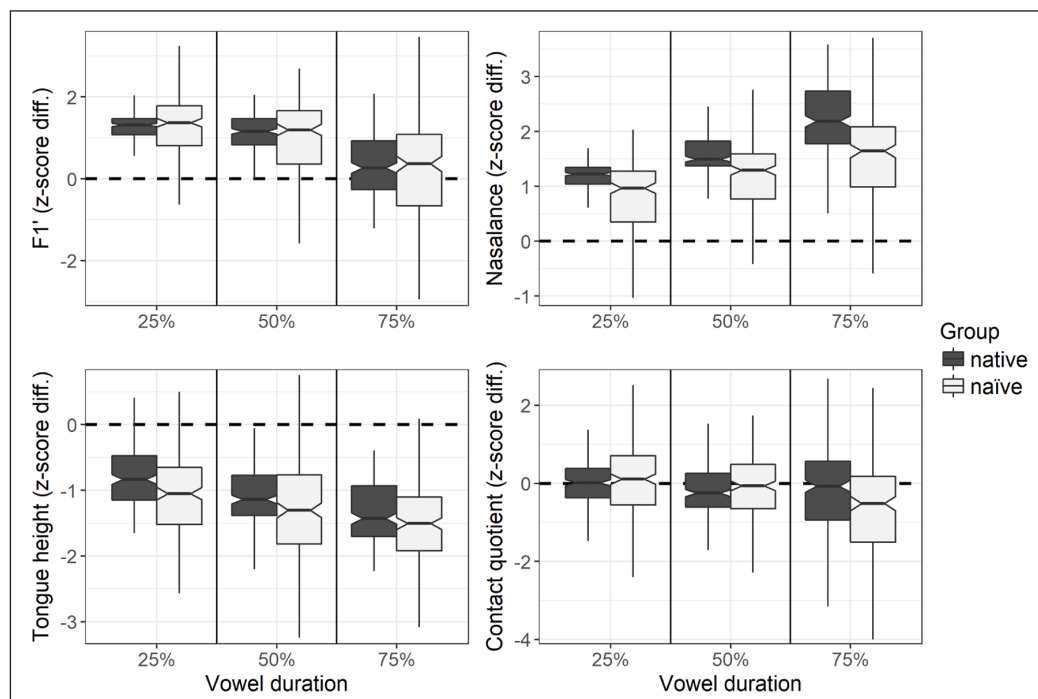
imitators and the native speakers. Taking CQ as an indicator of breathiness, these results suggest that / $\tilde{a}$ / was produced with more breathiness than / $a$ /, and that this articulatory difference was similar for the AE naïve imitators and the SF native speakers.

#### 4.1.2. / $\tilde{e}$ / vs. / $\epsilon$ / distinctions

Box plots of / $\tilde{e}$ /-/ $\epsilon$ / z-score differences for F1', nasalance, tongue height, and CQ values are shown in **Figure 4**. With regard to F1', the first LME model reveals that / $\tilde{e}$ / was produced with higher F1' compared to / $\epsilon$ /, but the box plots suggest that this distinction decreases throughout the vowel interval. No significant effect of GROUP was observed in the second LME model, revealing that the speaker-normalized F1' distinction between / $\tilde{e}$ / and / $\epsilon$ / was the same for both the naïve imitators and the native speakers. These results suggest that the AE naïve imitators were generally successful in replicating the F1'-frequency-related acoustic pattern produced by the SF native speakers.

With regard to nasalance, the same pattern can be observed for the vowel pair / $\tilde{e}$ /-/ $\epsilon$ / as was observed for the vowel pair / $\tilde{a}$ /-/ $a$ /: / $\tilde{e}$ / was produced with significantly more nasalance than / $\epsilon$ /, and the nasalance distinction increased in linear manner throughout the vowel interval. The significant (negative) effect of GROUP in the second LME model reveals that the AE imitators produced a smaller nasalance distinction between / $\tilde{e}$ / and / $\epsilon$ /, in comparison to the SF speakers.

With regard to tongue height, the first LME model reveals that / $\tilde{e}$ / was produced with a lower tongue position than / $\epsilon$ /, and the box plots suggest that this distinction increased slightly throughout the vowel interval. The significant (negative) effect of GROUP in the second LME model reveals that the AE imitators produced a larger tongue height distinction between / $\tilde{e}$ / and / $\epsilon$ /, in comparison to the SF speakers. In other words, while both groups produced / $\tilde{e}$ / with a lower tongue position than / $\epsilon$ /, this (negative) difference in tongue height between the two vowels was larger for the AE naïve imitators than for the SF native speakers.



**Figure 4:** / $\tilde{e}$ /-/ $\epsilon$ / z-score differences for F1', nasalance, tongue height, and CQ, as produced by the SF native speakers (dark box plots) and the AE naïve imitators (light box plots). Separate box plots are displayed for measurements taken at 25%, 50%, and 75% of the vowel duration.

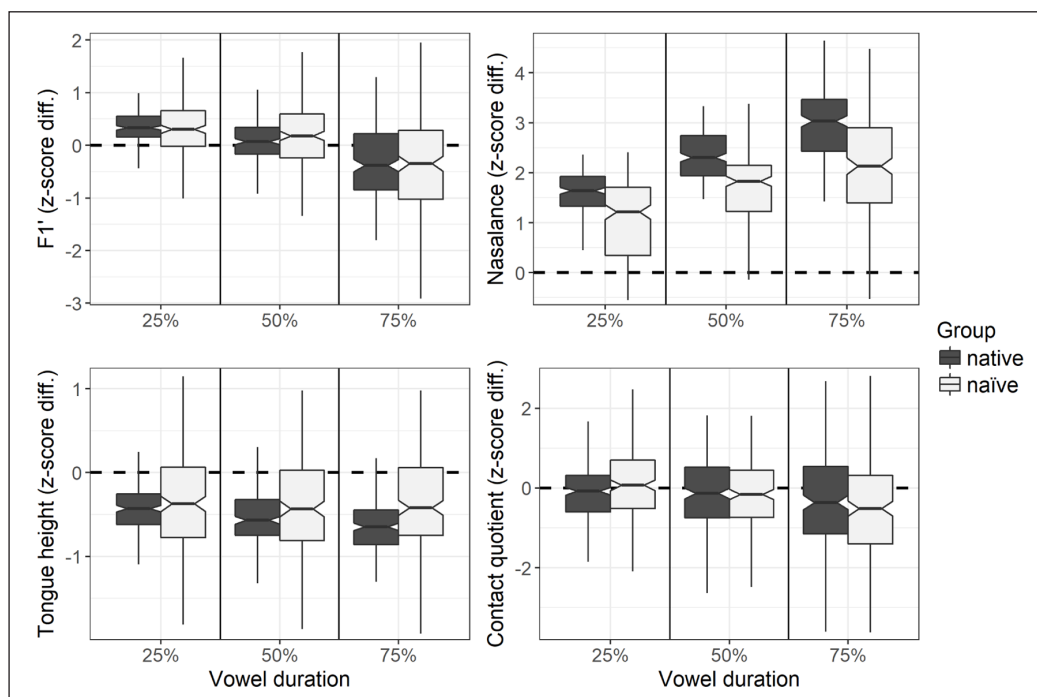


With regard to CQ, the first LME model reveals that / $\bar{\epsilon}$ / was produced with lower CQ values than / $\epsilon$ /, although the box plots suggest that this difference is slight. No significant effect of GROUP was observed in the second LME model, revealing that the speaker-normalized CQ distinction between / $\bar{\epsilon}$ / and / $\epsilon$ / was the same for both groups. Taking CQ as an indicator of breathiness, these results suggest that / $\bar{\epsilon}$ / was produced with more breathiness than / $\epsilon$ /, and that this articulatory difference was similar for the AE naïve imitators and the SF native speakers.

#### 4.1.3. / $\bar{\imath}$ / vs. / $\imath$ / distinctions

Box plots of / $\bar{\imath}$ /-/ $\imath$ / z-score differences for F1', nasalance, tongue height, and CQ values are shown in **Figure 5**. With regard to F1', the first LME model reveals that / $\bar{\imath}$ / was produced with no overall difference in F1' compared to / $\imath$ /. However, the box plots suggest that the dynamic nature of the F1' distinction is the likely cause for the overall non-significant result: In comparison to / $\imath$ /, / $\bar{\imath}$ / was produced with higher F1' values in the first part of the vowel interval, no difference at the vowel midpoint, and lower F1' values in the latter part of the vowel interval. As with / $\bar{\alpha}$ /-/ $\alpha$ / and / $\bar{\epsilon}$ /-/ $\epsilon$ /, this pattern also indicates a decline in frequency throughout the vowel interval; in the case of / $\bar{\imath}$ /-/ $\imath$ /, however, the pattern results in no overall F1' distinction between the two vowels. No significant effect of GROUP was observed in the second LME model, revealing that the speaker-normalized F1' distinction between / $\bar{\imath}$ / and / $\imath$ / was the same for both the naïve imitators and the native speakers. These results suggest that the AE naïve imitators were generally successful in replicating the F1'-frequency-related acoustic pattern produced by the SF native speakers, as was also observed for / $\bar{\alpha}$ /-/ $\alpha$ / and / $\bar{\epsilon}$ /-/ $\epsilon$ /.

Although no overall F1' distinction was observed between / $\bar{\imath}$ / and / $\imath$ /, there were significant differences in the articulatory parameters for the two vowels. With regard to nasalance, the first LME model and the box plots reveal the same pattern as was observed for / $\bar{\alpha}$ /-/ $\alpha$ / and / $\bar{\epsilon}$ /-/ $\epsilon$ /: / $\bar{\imath}$ / was produced with more nasalance than / $\imath$ /, and



**Figure 5:** / $\bar{\imath}$ /-/ $\imath$ / z-score differences for F1', nasalance, tongue height, and CQ, as produced by the SF native speakers (dark box plots) and the AE naïve imitators (light box plots). Separate box plots are displayed for measurements taken at 25%, 50%, and 75% of the vowel duration.

the nasalance distinction increases in linear manner throughout the vowel interval. The significant (negative) effect of GROUP in the second LME model reveals that the AE imitators produced a smaller nasalance distinction between /ĩ/ and /o/, in comparison to the SF speakers.

With regard to tongue height, although the two vowels were realized with similar F1' values, the first LME model reveals that /ĩ/ was produced with a significantly lower tongue position than /o/. The significant (positive) effect of GROUP in the second LME model reveals that the AE imitators produced a smaller tongue height distinction between /ĩ/ and /o/, in comparison to the SF speakers. In other words, while both groups produced /ĩ/ with a lower tongue position than /o/, this (negative) difference in tongue height between the two vowels was not as large for the AE naïve imitators as for the SF native speakers.

With regard to CQ, the first LME model reveals that /ĩ/ was produced with lower CQ values compared to /o/. However, no significant effect of GROUP was observed in the second LME model, revealing that the speaker-normalized CQ distinction between /ĩ/ and /o/ was the same for both the naïve imitators and the native speakers. Taking CQ as an indicator of breathiness, these results suggest that /ĩ/ was produced with more breathiness than /o/, and that this articulatory difference was similar for the AE naïve imitators and the SF native speakers.

#### 4.1.4. Summary: Articulatory and F1' nasal-oral distinctions

In summary, no significant effects of GROUP on the F1' nasal-oral distinction were observed for any of the vowel pairs, which suggests that the AE naïve imitators produced a similar acoustic output to the SF native speaker realizations, at least with regard to F1' frequency. Nevertheless, GROUP was found to be a significant factor in both the nasalance distinction and the tongue height distinction for all three vowel pairs. With regard to nasalance, the AE imitators successfully produced all three nasal vowels with a greater degree of nasalance throughout the entire vowel interval, in comparison to their oral vowel counterparts; however, this articulatory distinction was smaller in each of the three nasal-oral vowel pairs, in comparison to what the SF speakers had originally produced. With regard to tongue height, different patterns emerged for the three vowel pairs: The SF native speakers produced /ã/ with a lower tongue position than /a/, while the AE naïve imitators produced /ã/ with a higher tongue position than /a/; both groups produced /ẽ/ with a lower tongue position than /ɛ/, but this difference was even greater for the AE imitators than for the SF speakers; both groups produced /ĩ/ with a lower tongue position than /o/, but this difference was not as large for the AE imitators as for the SF speakers. Finally, all three nasal vowels were produced with lower CQ values compared to their oral vowel counterparts, and no significant differences in this CQ distinction were observed between the AE imitators and the SF speakers for any of the three vowel pairs. In other words, all three nasal vowels were produced with greater breathiness than their oral congeners, and this articulatory distinction was the same for the two groups.

#### **4.2. Relationship between articulatory distinctions and F1' distinction**

Since no difference between the naïve imitators and the native speakers was observed for the F1' distinction in any of the three nasal-oral vowel pairs, but significant differences were observed for both the nasalance distinction and the tongue height distinction for all three vowel pairs, we now turn to whether the relationship between the articulatory distinctions and the F1' distinction is different between the two groups. The results from the LME models constructed to investigate these possible differences are detailed in this section; the complete model results are provided in Table E of Appendix A.

#### 4.2.1. F1' ~ nasalance

The results from the LME models constructed to test the effect of NASALANCE and the interaction between NASALANCE and GROUP on F1' are the same for all three vowel pairs. Firstly, a significant (negative) main effect of NASALANCE is observed for all three vowel pairs: As the nasal vowels are produced with increasingly greater nasalance than the oral vowels, F1' frequency decreases systematically. This result is in agreement with the predicted effect of nasalization on F1' for the three nasal vowels studied here, all of which are non-high vowels (see Section 1.3.1). Secondly, a significant (positive) interaction between NASALANCE and GROUP is observed for all three vowel pairs. This result reveals that, although there is a negative relationship between the nasalance distinction and the F1' distinction for all three vowel pairs, this negative relationship is less strong for the AE naïve speakers compared to the SF native speakers in each of the three cases.

#### 4.2.2. F1' ~ tongue height

The results from the LME models constructed to test the effect of TONGUE HEIGHT and the interaction between TONGUE HEIGHT and GROUP on F1' are somewhat less straightforward than for the effect of NASALANCE. With regard to the main effect of TONGUE HEIGHT, no significant differences are observed for either /ã/-/a/ or /õ/-/o/. This result suggests that differences in tongue height between the respective nasal and oral congeners in these two vowel pairs do not systematically predict differences in F1'. However, a significant (positive) main effect of TONGUE HEIGHT is observed for the vowel pair /ẽ/-/ɛ/. Since it was observed in the previous section that /ẽ/ was produced with a lower tongue position than /ɛ/, this result indicates that as the tongue height distinction between the two vowels decreases (i.e., the tongue position for /ẽ/ increases towards that of /ɛ/), F1' increases systematically. Conversely, as the tongue height distinction between the two vowels increases (i.e., the tongue position for /ẽ/ decreases relative to /ɛ/), F1' decreases systematically. This result is not in agreement with the general prediction that tongue height is inversely correlated with F1 (Johnson, 2003; Stevens, 2000). Although the global distinction for this vowel pair displays the predicted pattern (i.e., /ẽ/ is produced with an overall lower tongue position, and is realized with an overall higher F1', compared to /ɛ/), and the overall correlation between tongue height and F1' for the three vowel pairs combined is indeed negative ( $R = -0.42, p < 0.001$ ), the absence of this relationship for /ã/-/a/ and /õ/-/o/, as well as the observation of a positive relationship for /ẽ/-/ɛ/, challenges this general assumption when considering within-vowel variation. However, given the significant relationships between nasalization and F1' (see above) and breathiness and F1' (see below), the absence of the expected within-vowel-pair relationship between tongue height and F1' is likely due to the influence of these articulatory variables.

A more consistent pattern can be seen for the interaction between TONGUE HEIGHT and GROUP on the realization of F1' frequency: A significant (negative) interaction in the LME models is observed for all three vowel pairs. This suggests that, while an overall pattern of an inverse relationship between tongue height and F1' is not observed, the AE naïve imitators display evidence of this very pattern when compared with the SF native speakers. In other words, as the nasal vowels are produced with an increasingly higher tongue position relative to the oral vowels, F1' frequency difference decreases systematically; or, conversely, as the nasal vowels are produced with increasingly lower tongue position relative to the oral vowels, F1' frequency difference increases systematically. However, this pattern is only observed for the AE imitators or is, at the very least, stronger for the AE imitators compared to the SF speakers.

#### 4.2.3. F1' ~ contact quotient

The results from the LME models constructed to test the effect of CQ and the interaction between CQ and GROUP on F1' reveal a significant (positive) main effect of CQ for all three vowel pairs: As the nasal vowels are produced with increasingly greater CQ values, F1' frequency increases systematically. Given the results from the previous section, the converse interpretation is perhaps more meaningful: As the nasal vowels are produced with decreasing CQ values (i.e., as breathiness increases), F1' decreases systematically. This result is in agreement with the predicted effect of breathy voicing on F1' frequency for the non-high vowels studied here (see Section 1.3.1). With regard to the interaction between CQ and GROUP, the only significant effect is observed for the vowel pair /ã/-/a/: Although no differences in CQ were observed between the two groups (see Section 4.2), the positive relationship between CQ and F1' is not as strong for the AE imitators compared to the SF speakers for this vowel pair.

## 5. Discussion

### 5.1. Summary and discussion of results

**Table 1** provides a summary of the results for the F1' and articulatory nasal-oral distinctions (Nasality), the differences between the AE naïve imitators and the SF native speakers with regard to these distinctions (Group), and the relationship between the articulatory variables and F1' as well as the differences between the AE imitators and the SF speakers with regard to these relationships (F1' ~ IV). For the Nasality results, up and down arrows correspond to whether the given variable was significantly higher

**Table 1:** Summary of results. Nasality: Symbols denote whether the nasal vowel was higher than (↑), lower than (↓), or equal to (=) its oral counterpart. Group: Relationships denote whether the nasal-oral distinction was higher ( $AE > SF$ ), lower ( $AE < SF$ ), or equal ( $AE = SF$ ) for the AE imitators vs. the SF speakers. F1' ~ IV: Symbols denote whether the articulatory variable positively (+) or negatively (-) predicted F1'; relationships denote whether this articulatory-F1' mapping was more positive ( $AE > SF$ ), more negative ( $AE < SF$ ), or equal ( $AE = SF$ ) for the AE imitators vs. the SF speakers.

Variable	Comparison	Vowel pair		
		/ã/ - /a/	/ẽ/ - /ɛ/	/õ/ - /o/
F1'	Nasality	↓	↑	=
	Group	$AE = SF$	$AE = SF$	$AE = SF$
	F1' ~ IV	NA	NA	NA
Nasalance	Nasality	↑	↑	↑
	Group	$AE < SF$	$AE < SF$	$AE < SF$
	F1' ~ IV	- $AE > SF$	- $AE > SF$	- $AE > SF$
Tongue height	Nasality	=	↓	↓
	Group	$AE > SF$	$AE < SF$	$AE > SF$
	F1' ~ IV	$AE < SF$	+ $AE < SF$	$AE < SF$
Breathiness	Nasality	↓	↓	↓
	Group	$AE = SF$	$AE = SF$	$AE = SF$
	F1' ~ IV	+ $AE < SF$	+ $AE = SF$	+ $AE = SF$

(↑) or lower (↓) for the nasal vowel compared to its oral counterpart; differences that were not significant are denoted as equal. For the Group results, nasal-oral distinctions that were higher (i.e., positive distinctions that were *more positive* and negative distinctions that were *less negative*) for the AE imitators vs. the SF speakers are denoted by  $AE > SF$ ; nasal-oral distinctions that were lower (i.e., positive distinctions that were *less positive* and negative distinctions that were *more negative*) for the AE imitators vs. the SF speakers are denoted by  $AE < SF$ ; nasal-oral distinctions that were the same for both groups are denoted by  $AE = SF$ . For the  $F1' \sim IV$  results, articulatory variables that positively predicted  $F1'$  are denoted by +, and articulatory variables that negatively predicted  $F1'$  are denoted by -; articulatory- $F1'$  mappings that yielded a more positive relationship for the AE imitators vs. the SF speakers are denoted by  $AE > SF$ ; articulatory- $F1'$  mappings that yielded a more negative relationship for the AE imitators vs. the SF speakers are denoted by  $AE < SF$ ; articulatory- $F1'$  mappings that yielded the same relationship between the two groups are denoted by  $AE = SF$ . By way of example, we take the nasalance results for the vowel pair /ã/-/a/: Nasalance was greater for /ã/ vs. /a/ (↑), but this distinction was smaller for the AE imitators vs. the SF speakers ( $AE < SF$ ); nasalance negatively predicted  $F1'$  (-), but this relationship was less strong/negative (i.e., more positive slope) for the AE imitators vs. the SF speakers ( $AE > SF$ ).

With regard to  $F1'$ , /ã/ was realized with lower  $F1'$  values than /a/, /ẽ/ was realized with higher  $F1'$  values than /ε/, and /õ/ and /o/ were realized with the same overall  $F1'$  values. The naïve imitators reproduced the same nasal-oral distinctions as the native speakers for every vowel pair. This can be taken as an indication of the imitators faithfully replicating the acoustic output produced by the native speakers—at least with regard to the  $F1'$  frequency distinction, which was the specific focus of the current study—and, thus, successfully completing the experimental task. However, the results for the articulatory variables suggest that the imitators sometimes reached these acoustic goals via different articulatory strategies than those used by the native speakers. Moreover, differences between the AE imitators and the SF speakers were observed for the relationship between the articulatory variables and  $F1'$  frequency.

With regard to nasalance, the AE imitators produced the same distinctions as the SF native speakers: greater nasalance for all of the nasal vowels compared to their oral counterparts. However, this distinction was smaller for the imitators compared to the native speakers in every case. It could be argued that the difference in magnitude between the two groups was due to the AE imitators having had no exposure to a language with phonemic vowel nasality and, thus, being unable to produce distinctive nasality as part of their productions. However, the fact that the naïve imitators reliably produced a distinction between the oral and nasal vowels throughout the entire vowel interval (at least, from 25% of the vowel interval, as measured in this study) suggests that this is not the case: Naïve speakers of a language without contrastive vowel nasality are indeed able to produce contrastive vowel nasality when imitating contrastive nasal vowels. With regard to the relationship between nasalance and  $F1'$ , nasalance negatively predicted  $F1'$  frequency for all three vowel pairs, as expected for all of the non-high vowels in this study (see Section 1.3.1). However, this relationship was less strong for the AE imitators compared to the SF speakers in every case (i.e., more positive slope for the AE vs. SF GROUP interaction). In other words, while the AE imitators produced a nasal-oral distinction using degree of nasalization, the distinction was smaller for the imitators compared to the native speakers, and the lowering effect on  $F1'$  was less apparent. These results suggest that, while both groups used degree of nasalization in a way that affected  $F1'$  frequency for all three nasal-oral vowel pairs, the

AE naïve imitators employed this articulatory-acoustic relationship to a lesser degree than the SF native speakers had originally used.

With regard to tongue height, the results were different for the three vowel pairs. For /ã/-/a/, although there was no overall difference in height between the two vowels, this aggregate result was likely due to different patterns for the two groups: The SF native speakers produced /ã/ with a *lower* tongue position than /a/, overall, while the AE naïve imitators produced /ã/ with a *higher* tongue position than /a/, overall. Indeed, the LME model tested on the randomized matched nasal-oral comparisons revealed that the AE imitators produced a more positive /ã/-/a/ tongue height distinction compared to the SF speakers. Moreover, separate post-hoc LME models revealed that the two group-wise patterns (i.e., lower tongue position for /ã/ vs. /a/ for the SF speakers and higher tongue position for /ã/ vs. /a/ for the AE imitators) were both significant. For /ẽ/-/ε/, /ẽ/ was produced with a lower tongue position than /ε/, overall. However, this distinction was even greater for the AE imitators than for the SF speakers: The imitators produced /ẽ/ with an even lower tongue position vs. /ε/, in comparison to the native speakers. For /õ/-/o/, /õ/ was produced with a lower tongue position than /o/, overall. However, this distinction was not as great for the AE imitators compared to the SF speakers. With regard to the relationship between tongue height and F1', although tongue height negatively predicted F1' frequency when the data for all three vowel pairs were included together (i.e., the expected inverse relationship between tongue height and F1), this relationship was not observed for any of the three nasal-oral vowel pairs, individually. In fact, an overall positive relationship between tongue height and F1' was observed for the vowel pair /ẽ/-/ε/. However, in comparing the AE imitator tongue height/F1' relationships to those produced by the SF speakers, the AE imitators had a significantly more negative relationship in every case. In other words, although the expected inverse relationship between tongue height and F1' was not observed for individual vowel pairs, this inverse relationship was revealed in the difference between the AE imitators and the SF speakers for each of the three vowel pairs. These results suggest that the AE naïve imitators employed this articulatory-acoustic relationship to a greater degree than the SF native speakers had originally used.

With regard to breathiness, the AE imitators produced the same distinctions as the SF native speakers: lower CQ values (i.e., greater breathiness) for all of the nasal vowels compared to their oral vowel counterparts. Furthermore, the AE imitators produced this distinction to the same degree as the SF native speakers in every case. With regard to the relationship between breathiness and F1', CQ positively predicted F1' frequency for all three vowel pairs. In other words, since CQ is inversely related to breathiness, breathiness negatively predicted F1' frequency for all three vowel pairs, as predicted for all of the non-high vowels in this study (see Section 1.3.1). This articulatory-acoustic relationship was the same between the two groups for the vowel pairs /ẽ/-/ε/ and /õ/-/o/, but the relationship was less strong for the AE imitators vs. the SF speakers for the vowel pair /ã/-/a/. These results suggest that both groups used breathiness in a way that affects F1' frequency for all three nasal-oral vowel pairs, and that the AE naïve imitators employed this articulatory-acoustic relationship to the same degree as the SF native speakers for two of the three vowel pairs, but to a lesser degree than the SF speakers had originally used for /ã/-/a/.

## 5.2. Interpretations for mechanisms of listener-based sound change

The hypotheses for the two theoretical frameworks for listener-based sound change detailed in Section 1.2 were as follows:

1. Evidence in support of listener misperception (Ohala):
  - a. The SF speakers produced articulatory distinctions that the AE imitators fail to replicate.
  - b. The AE imitators produce articulatory distinctions that the SF speakers had not originally produced.
2. Evidence in support of perceptual cue re-weighting (Beddor):

The AE imitators produce the same articulatory distinctions as the SF speakers but to different degrees.

No direct evidence for hypotheses 1a or 1b, supporting listener misperception (Ohala, 1981, 1993), was observed in the current study: Of the three articulatory variables that were monitored in this study (degree of nasalization, tongue height, breathiness) no cases were observed of the imitators employing an articulator that was never originally used by the native speakers, or vice versa.<sup>4</sup> However, the results for tongue height and F1' for the vowel pair /ã/-/a/ can be argued to support a misperception-based account in a manner that synthesizes these two hypotheses. Although the SF native speakers produced /ã/ with a lower tongue position than /a/ (predicted to raise F1' for /ã/ vs. /a/), /ã/ was nonetheless realized with a lower F1' frequency than /a/. Given that /ã/ was found to be produced by the SF speakers with greater degrees of nasalization and breathiness compared to /a/, and that these articulatory distinctions were found to be significantly predictive of lower F1' values, it is likely that the lower F1' frequency observed for /ã/ vs. /a/ arose due to these articulatory distinctions, and not to tongue height. In comparison, although the AE naïve imitators produced the same F1' distinction that the SF speakers had produced (and to the same degree), they did so by using a different articulatory strategy than the SF speakers had used. In particular, the AE imitators produced /ã/ with a *higher* tongue position than /a/, which is expected to lower F1' frequency. Thus, these results suggest that the SF native speakers had produced an articulatory distinction that the AE naïve imitators failed to replicate (i.e., a lower tongue position for /ã/ vs. /a/), while the AE imitators produced instead an articulatory distinction that the SF speakers had not originally made (i.e., a higher tongue position for /ã/ vs. /a/). This result indicates that the AE imitators may have misperceived the lower F1' frequency for /ã/ vs. /a/ as being due to a higher tongue position, and subsequently produced this articulatory distinction in their imitations.

Evidence for hypothesis 2, supporting perceptual cue re-weighting of co-varying articulatory properties (Beddor, 2009, 2012), was observed in the current study. This evidence was observed primarily in the differences between the SF speakers and the AE imitators with regard to how nasalance and tongue height related to F1' frequency for the two groups. Crucially, the AE imitators successfully produced the same F1' nasal-oral distinctions that the SF speakers had originally produced for each of the three vowel pairs, and they produced these distinctions to the same degree. However, the articulatory evidence suggests that the AE imitators reached these F1' targets using a different weighting of the various articulatory distinctions compared to the SF native speakers. On the one hand, both groups used breathiness in the same manner (i.e., all nasal vowels were breathier than their oral counterparts), and increased breathiness was predictive of

---

<sup>4</sup> It should be noted that hypotheses 1a and 1b represent a particularly narrow interpretation of the Ohalan framework of misperception. The overall results from the current study may, in fact, be compatible with a broader interpretation than what has been presented here: The imitators may have correctly detected the presence of nasality in the signal but misperceived the degree to which nasality contributed to F1' modulation, concurrently misperceiving the contribution of tongue height to F1' modulation as well.

decreased F1' for both groups (although this relation was less strong for the AE imitators' productions of /ã/-/a/ compared to the SF speakers). On the other hand, evidence was observed of a trading relation between nasalance and tongue height with regard to how the two groups employed these articulatory variables to affect F1' frequency. Overall, an increase in nasalance was found to significantly predict a decrease in F1' frequency; however, the AE imitators produced a smaller nasal-oral nasalance distinction for all of the vowel pairs and, accordingly, the relationship between nasalance and F1' was less strong for the AE imitators in comparison to the SF speakers for all three vowel pairs. Increased tongue height was found to be inversely correlated with F1' for all of the vowel data combined (although this relationship was not observed for any of the individual vowel pairs); however, this inverse relationship between tongue height and F1' was stronger for the AE imitators in comparison to the SF speakers for all three vowel pairs. Taken together, these results suggest that, in comparison to the SF native speakers, the AE naïve imitators employed nasalance less, and tongue height more, to affect F1' frequency in their nasal-oral vowel distinctions. Since the F1' frequency distinctions were the same for the two groups, this indicates that the AE imitators may have re-weighted the relative magnitudes of the nasalance and tongue height distinctions in the SF speaker productions, assigning different importance to these variables (i.e., less to nasalance and more to tongue height) in order to obtain a similar F1' frequency output as the SF speakers.

An alternative explanation for the trading relation between the tongue height/F1' relationship and the nasalization/F1' relationship concerns the respective phonological representations of these two articulations in French and English. When a phonetic effect (such as nasalization) becomes phonologized in a given language, the articulatory distinction is sometimes enhanced beyond what is observed in purely phonetic variation (Hyman, 2013; Solé, 2007). Thus, since vowel nasalization is phonemic in (Southern) French, but phonetic in (Australian) English, the observation that the AE imitators produced a smaller nasalance distinction than the SF speakers can be explained by the different phonological representations of vowel nasality in the two languages.<sup>5</sup> Likewise, the fact that the AE imitators realized a stronger inverse relationship between tongue height and F1' could be explained by their native phonology—more precisely, since neither vowel nasality nor breathy voice are phonologically contrastive in English, the imitators may have relied more strongly on tongue height in order to realize F1' distinctions between oral and nasal vowel counterparts. However, there are two observations from this study that provide possible counter-evidence to this interpretation. Firstly, the AE imitators (re-)produced the same F1' nasal-oral distinctions that the SF speakers had originally produced for each of the three vowel pairs. If the trading relation between nasalance and tongue height was due simply to differences in the phonological representations of these articulations in the two languages, there is no a priori reason to expect that the acoustic realizations should be the same for the two groups; rather, the results from this study suggest that the AE imitators re-weighted these two articulatory variables in order to achieve the same acoustic output (here, F1' frequency) as the SF speakers. Secondly, breathy voicing is not phonologically contrastive for vowels in either French or English; nonetheless, both groups produced all three nasal vowels with a greater degree of breathiness compared to their oral counterparts. If the phonetic differences in nasalization and tongue height between the two groups were due simply to differences in the phonological representation of these articulations in the two languages, then we should not necessarily expect either group to produce phonetically distinctive *breathiness* as part of the nasal-oral contrast, yet

<sup>5</sup> I would like to thank Jonathan Harrington for this particular interpretation of the results.



this is what has been observed here. Ultimately, the trading relation between the tongue height/F1' relationship and the nasalization/F1' relationship may also be compatible with a hybrid account of cue re-weighting and language-specific phonology: English phonology may contribute to the AE imitators' preferred cue weighting, biasing their productions. In other words, it may not necessarily be the case that the AE imitators re-weighted the cues present in the speech signal in an independent, unconstrained manner. Rather, the cue weights of their native phonology—i.e., more weight given to the tongue height/F1' relationship and less weight given to the nasalization/F1' relationship—may have influenced how the re-weighting strategy was manifested in their imitations.<sup>6</sup>

## 6. Conclusion

The current study was designed to test whether evidence of mechanisms of listener-based sound change—listener misperception (Ohala, 1981, 1993) and perceptual cue re-weighting (Beddor, 2009, 2012)—could be observed in a laboratory setting. In order to elicit possible evidence of these mechanisms, articulatory and acoustic data related to the productions of phonemic oral and nasal vowels of Southern French (SF) were collected initially from four native speakers (Carignan, 2017). Acoustic data included measurements of F1' frequency; articulatory data included measurements of degree of nasalization, tongue height, and breathiness, due to the independent effects that these articulations have on F1' frequency, specifically. The acoustic data from these recordings were presented to nine Australian English (AE) naïve listeners, who were instructed to imitate the native SF productions. During these imitations, similar articulatory and acoustic data were collected in order to compare the naïve imitator articulatory strategies to those used by the native speakers.

In comparing the naïve AE imitations to the native SF productions, the results suggest that the AE imitators successfully reproduced the F1' frequency distinction made by the SF native speakers for all three vowel pairs (i.e., higher F1' for /*ẽ*/ vs. /*ɛ*/, lower F1' for /*ã*/ vs. /*a*/, no difference for /*õ*/ vs. /*o*/). However, differences were observed with regard to the articulatory strategies that the two groups employed to obtain these acoustic distinctions. The articulatory strategies for the vowel pair /*ã*/-/*a*/ suggest that listeners (at least partially) misperceived F1'-lowering due to nasalization and breathiness as being due to tongue height. Additional evidence supported perceptual cue re-weighting, particularly with regard to the relation between nasalance, tongue height, and F1' frequency. Specifically, evidence was observed that the AE imitators used nasalance less, and tongue height more, in order to obtain the same F1' nasal-oral distinctions that the SF speakers had originally produced.

Future research in this domain—involving different language backgrounds, different speech sounds, and/or different articulatory-to-acoustic relationships—would help to determine if the results observed here are idiosyncratic to the current study or, instead, part of a more general cognitive trait of the human speech faculty that fosters the emergence of linguistic diversity and sound change from interactions between speakers and hearers.

## Additional Files

The additional files for this article can be found as follows:

- **Appendix A.** Miscellaneous tables. DOI: <https://doi.org/10.5334/labphon.136.s1>
- **Appendix B.** Individual plots for native speakers and naïve imitators. DOI: <https://doi.org/10.5334/labphon.136.s2>

<sup>6</sup> I would like to thank Pam Beddor for this particular interpretation of the results.

## Acknowledgements

This research was funded by The MARCS Institute for Brain, Behaviour and Development at Western Sydney University and ERC Advanced Grant 295573 “Human interaction and the evolution of spoken accent” (J. Harrington). Portions of the research have previously appeared in the *Journal of Phonetics* (Carignan, 2017) and *Journal of the Acoustical Society of America* (Carignan, 2018). The author gratefully acknowledges Ander Egurtzegi for his comments on an earlier draft of this manuscript, as well as Jonathan Harrington, Pam Beddor, and an anonymous reviewer for their careful and attentive examination of this work. The author is also thankful for comments and suggestions from attendees at the *Workshop on the Regularity Principle in Sound Change* and the *3<sup>rd</sup> International Workshop on Dynamic Modeling*, which both took place at University of Cologne in July, 2017, and from attendees of *LabPhon16*, which took place at Universidade de Lisboa in June, 2018.

## Competing Interests

The author has no competing interests to declare.

## References

- Arai, T. 2004. Comparing tongue positions of vowels in oral and nasal contexts. In: *Sophia Symposium on New Technology and the Investigation of the Articulatory Process*, 33–49. Sophia University.
- Audacity Development Team. 2016. *Audacity 2.1.2*. Computer software program available from: <https://www.audacityteam.org>.
- Baken, R. J., & Orlikoff, R. F. 2000. *Clinical Measurement of Speech and Voice*, 2nd ed. San Diego, CA: Singular Publishing Group.
- Barlaz, M., Fu, M., Dubin, J., Shosted, R., Liang, Z.-P., & Sutton, B. 2015. Lingual differences in Brazilian Portuguese: An rt-MRI study. In: *Proceedings of the 18th Meeting of the International Congress of the Phonetic Sciences*.
- Bates, D., Mächler, M., Bolker, B., & Walker, S. 2015. Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software*, 67(1), 1–48. DOI: <https://doi.org/10.18637/jss.v067.i01>
- Beddor, P. S. 1982. *Phonological and phonetic effects of nasalization on vowel height* (Unpublished doctoral dissertation). University of Minnesota. Bloomington, IN: Indiana University Linguistics Club.
- Beddor, P. S. 2009. A coarticulatory path to sound change. *Language*, 85(4), 785–821. DOI: <https://doi.org/10.1353/lan.0.0165>
- Beddor, P. S. 2012. Perception grammars and sound change. In: Solé, M.-J., & Recasens, D. (eds.), *The initiation of sound change. Perception, production, and social factors*, 37–55. Amsterdam: John Benjamins. DOI: <https://doi.org/10.1075/cilt.323.06bed>
- Beddor, P. S., Krakow, R. A., & Goldstein, L. M. 1986. Perceptual constraints and phonological change: A study of nasal vowel height. *Phonology Yearbook*, 3, 197–217. DOI: <https://doi.org/10.1017/S0952675700000646>
- Bickley, C. 1982. Acoustic analysis and perception of breathy vowels. *MIT Speech Communication Working Papers*, 1, 71–81.
- Bloch, J. 1920. *La formation de la langue marathe*. Paris: Champion.
- Bloch, J. 1965. *Indo-Aryan: From the Vedas to modern times*. Paris: Maisonneuve.
- Boersma, P., & Weenink, D. 2015. *Praat: doing phonetics by computer*. Computer software program available from: <http://www.praat.org/>.
- Bothorel, A., Simon, P., Wioland, F., & Zerling, J.-P. 1986. Cinéradiographie des voyelles et consonnes du français. *Travaux de l'Institut de Phonétique de Strasbourg*, 18.

- Brichler-Labaeye, C. 1970. *Les voyelles françaises. Mouvements et positions articulatoires à la lumière de la radiocinématographie* (coll. Bibliothèque française et romane, série A, no. 18). Paris: Klincksieck.
- Carignan, C. 2013. *When nasal is more than nasal: The oral articulation of nasal vowels in two dialects of French*. (Unpublished doctoral dissertation). University of Illinois at Urbana-Champaign.
- Carignan, C. 2014. An acoustic and articulatory examination of the ‘oral’ in ‘nasal’: The oral articulations of French nasal vowels are not arbitrary. *Journal of Phonetics*, 46, 23–33. DOI: <https://doi.org/10.1016/j.wocn.2014.05.001>
- Carignan, C. 2017. Covariation of nasalization, tongue height, and breathiness in the realization of F1 of Southern French nasal vowels. *Journal of Phonetics*, 63, 87–105. DOI: <https://doi.org/10.1016/j.wocn.2017.04.005>
- Carignan, C. 2018. Using ultrasound and nasalance to separate oral and nasal contributions to formant frequencies of nasalized vowels. *Journal of the Acoustical Society of America*, 143(5), 2588–2601. DOI: <https://doi.org/10.1121/1.5034760>
- Carignan, C., Shosted, R., Fu, M., Z.-P., & Sutton, B. 2015. A real-time MRI investigation of the role of lingual and pharyngeal articulation in the production of the nasal vowel system of French. *Journal of Phonetics*, 50, 34–51. DOI: <https://doi.org/10.1016/j.wocn.2015.01.001>
- Carignan, C., Shosted, R., Shih, C., & Rong, P. 2011. Compensatory articulation in American English nasalized vowels. *Journal of Phonetics*, 39, 668–682. DOI: <https://doi.org/10.1016/j.wocn.2011.07.005>
- Chen, M. Y. 1997. Acoustic correlates of English and French nasalized vowels. *Journal of the Acoustical Society of America*, 102, 2360–2370. DOI: <https://doi.org/10.1121/1.419620>
- Cohn, A. C. 1990. *Phonetic and phonological rules of nasalization* (Unpublished doctoral dissertation). University of California, Los Angeles. Published as *UCLA Working Papers in Linguistics* 76.
- da Matta Machado, M. 1993. Fenômenos de nasalização vocálica em português: Estudo cine-radiográfico. *Cadernos de estudos lingüísticos*, 25, 113–127.
- Delvaux, V. 2012. *Les voyelles nasales du français: Aérodynamique, articulation, acoustique et perception*. Bruxelles: Peter Lang.
- Delvaux, V., Metens, T., & Soquet, A. 2002. French nasal vowels: Acoustic and articulatory properties. In: *Proceedings of the 7th International Conference on Spoken Language Processing, 1*, 53–56. Denver.
- Demolin, D., Delvaux, V., Metens, T., & Soquet, A. 2003. Determination of velum opening for French nasal vowels by magnetic resonance imaging. *Journal of Voice*, 17(4), 454–467. DOI: [https://doi.org/10.1067/S0892-1997\(03\)00036-5](https://doi.org/10.1067/S0892-1997(03)00036-5)
- Denning, K. 1989. *The diachronic development of phonological voice quality, with special reference to Dinka and the other Nilotic languages* (Unpublished doctoral dissertation). Stanford University.
- Derrick, D., Best, C. T., & Fiasson, R. 2015. Non-metallic ultrasound probe holder for cocollection and co-registration with EMA. In *Proceedings of 18th International Congress of Phonetic Sciences (ICPhS)*, 1–5.
- Engwall, O., Delvaux, V., & Metens, T. 2006. *Interspeaker variation in the articulation of nasal vowels*. Available online at: <http://www.speech.kth.se/prod/publications/files/1926.pdf>.
- Escudero, P., Boersma, P., Rauber, A. S., & Bion, R. A. H. 2009. A cross-dialect acoustic description of vowels: Brazilian and European Portuguese. *Journal of the Acoustical Society of America*, 126(3), 1379–1393. DOI: <https://doi.org/10.1121/1.3180321>

- FFmpeg Development Team. 2016. *FFmpeg*. Computer software program available from: <https://ffmpeg.org>.
- Garellek, M. 2014. Voice quality strengthening and glottalization. *Journal of Phonetics*, 45, 106–113. DOI: <https://doi.org/10.1016/j.wocn.2014.04.001>
- Garellek, M., & Keating, P. 2011. The acoustic consequences of phonation and tone interactions in Jalapa Mazatec. *Journal of the International Phonetic Association*, 41, 185–205. DOI: <https://doi.org/10.1017/S0025100311000193>
- Garellek, M., Ritchart, A., & Kuang, J. 2016. Breathy voice during nasality: A crosslinguistic study. *Journal of Phonetics*, 59, 110–121. DOI: <https://doi.org/10.1016/j.wocn.2016.09.001>
- Garellek, M., Samlan, R., Gerratt, B. R., & Kreiman, J. 2016. Modeling the voice source in terms of spectral slopes. *Journal of the Acoustical Society of America*, 139, 1404–1410. DOI: <https://doi.org/10.1121/1.4944474>
- Gordon, M., & Ladefoged, P. 2001. Phonation types: A cross-linguistic overview. *Journal of Phonetics*, 29, 383–406. DOI: <https://doi.org/10.1006/jpho.2001.0147>
- Grierson, G. 1922. Spontaneous nasalization in the Indo-Aryan languages. *Journal of the Royal Asiatic Society of Great Britain*, 381–388.
- Guion, S. 1998. The role of perception in the sound change of velar palatalization. *Phonetica*, 55, 18–52. DOI: <https://doi.org/10.1159/000028423>
- Guthrie, M. 1967–1970. *Comparative Bantu*, 4. Gregg.
- Hajek, J. 1997. *Universals of sound change in nasalization*. Oxford: Blackwell.
- Hanson, H. M., K. N. Stevens, K. N., Kuo, H.-K. J., Chen, M. Y., & Slifka, J. 2001. Towards models of phonation. *Journal of Phonetics*, 29, 451–480. DOI: <https://doi.org/10.1006/jpho.2001.0146>
- Harrington, J., Kleber, F., Reubold, U., Schiel, F., & Stevens, M. In press. The phonetic basis of the origin and spread of sound change. In: Katz, W., & Assmann, P. (eds.), *The Routledge Handbook of Phonetics*.
- Hyman, L. M. 2013. Enlarging the scope of phonologization. In: Yu, A. (ed.), *Origins of Sound Change: Approaches to Phonologization*, 3–28. Oxford: Oxford University Press.
- Imatomi, S. 2005. Effects of breathy voice source on ratings of hypernasality. *The Cleft Palate – Craniofacial Journal*, 42(6), 641–648. DOI: <https://doi.org/10.1597/03-146.1>
- Johnson, K. 2003. *Acoustic and auditory phonetics*. Oxford: Blackwell.
- Kingston, J., & Macmillan, N. A. 1995. Integrality of nasalization and F1 in vowels in isolation and before oral and nasal consonants: A detection-theoretic application of the Garner paradigm. *Journal of the Acoustical Society of America*, 97, 1261–1285. DOI: <https://doi.org/10.1121/1.412169>
- Krakow, R. A., Beddor, P. S., & Goldstein, L. M. 1988. Coarticulatory influences on the perceived height of nasal vowels. *Journal of the Acoustical Society of America*, 83(3), 1146–1158. DOI: <https://doi.org/10.1121/1.396059>
- Kreiman, J., Shue, Y.-L., Chen, G., Iseli, M., Gerratt, B. R., Neubauer, J., & Alwan, A. 2012. Variability in the relationships among voice quality, harmonic amplitudes, open quotient, and glottal area waveform shape in sustained phonation. *Journal of the Acoustical Society of America*, 132, 2625–2632. DOI: <https://doi.org/10.1121/1.4747007>
- Kuehn, D. P., & Azzam, N. A. 1978. Anatomical characteristics of palatoglossus and the anterior faucial pillar. *Cleft Palate Journal*, 15(4), 349–359.
- Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. 2016. *lmerTest: Tests in Linear Mixed Effects Models*. Computer software program available from: <https://cran.rproject.org/package=lmerTest>.

- Li, M., Kambhamettu, C., & Stone, M. 2005. Automatic contour tracking in ultrasound images. *Clinical Linguistics and Phonetics*, 19(6–7), 545–554. DOI: <https://doi.org/10.1080/02699200500113616>
- Lotto, A. J., Holt, L. L., & Kluender, K. R. 1997. Effect of voice quality on perceived height of English vowels. *Phonetica*, 54, 76–93. DOI: <https://doi.org/10.1159/000262212>
- Maeda, S. 1993. Acoustics of vowel nasalization and articulatory shifts in French nasal vowels. In: Huffman, M. K., & Krakow, R. A. (eds.), *Nasals, Nasalization, and the Velum*, 5, 147–170. San Diego: Academic Press. DOI: <https://doi.org/10.1016/B978-0-12-360380-7.50010-7>
- Mann, V. A., & Repp, B. H. 1980. Influence of vocalic context on perception of the [S]-[s] distinction. *Perception and Psychophysics*, 28, 548–558. DOI: <https://doi.org/10.3758/BF03204377>
- Martins, P., Oliveira, C., Silva, S., & Teixeira, A. 2012. Velar movement in European Portuguese nasal vowels. In: *Proc. IberSpeech—VII Jornadas en Tecnología del Habla and II Iberian SLTech Workshop, Madrid, Spain*, 231–240.
- Matisoff, J. A. 1970. Glottal dissimilation and the Lahu high-rising tone: A tonogenetic case-study. *Journal of the American Oriental Society*, 90(1), 13–44. DOI: <https://doi.org/10.2307/598429>
- Matisoff, J. A. 1975. Rhinoglottophilia: The mysterious connection between nasality and glottality. In: Ferguson, C. A., Hyman, L. M., & Ohala, J. J. (eds.), *Nasálfest: Papers from a Symposium on Nasals and Nasalization*, 265–287. Palo Alto, CA: Stanford University Language Universals Project.
- Meillet, A. 1967. *The comparative method in historical linguistics*. Paris: Librairie Honoré Champion.
- Moisik, S. R., Lin, H., & Esling, J. H. 2014. A study of laryngeal gestures in Mandarin citation tones using simultaneous laryngoscopy and laryngeal ultrasound (SLLUS). *Journal of the International Phonetic Association*, 44, 21–58. DOI: <https://doi.org/10.1017/S0025100313000327>
- Ohala, J. J. 1975. Phonetic explanations for nasal sound patterns. In: Ferguson, C. A., Hyman, L. M., & Ohala, J. J. (eds.), *Nasálfest: Papers from a Symposium on Nasals and Nasalization*, 289–316. Palo Alto, CA: Stanford University Language Universals Project.
- Ohala, J. J. 1981. The listener as a source of sound change. In: Masek, C. S., Hendrick, R. A., & Miller, M. F. (eds.), *Papers from the Parasession on Language and Behavior*, 178–203. Chicago: Chicago Linguistics Society.
- Ohala, J. J. 1989. Sound change is drawn from a pool of synchronic variation. In: Breivik, L. E., & Jahr, E. H. (eds.), *Language Change: Contributions to the study of its causes*, 173–198. Berlin: Mouton de Gruyter.
- Ohala, J. J. 1993. The phonetics of sound change. In: Jones, C. (ed.), *Historical linguistics: problems and perspectives*, 237–278. London: Longman.
- Ohala, J. J., & Amador, M. 1981. Spontaneous nasalization. *The Journal of the Acoustical Society of America*, 69, S54. (Abstract). DOI: <https://doi.org/10.1121/1.386212>
- Ohala, J. J., & Busa, M. G. 1995. Nasal loss before voiceless fricatives: A perceptually-based sound change. *Rivista di Linguistica*, 7, 125–144.
- Ohala, J. J., & Ohala, M. 1992. Nasals and nasalization in Hindi. *Paper presented at 3rd International Symposium on Language and Linguistics: Pan-Asiatic Linguistics*. Chulalongkorn Univ., Bangkok, Thailand. 8–10 Jan 92.
- Ohala, J. J., & Ohala, M. 1993. The phonetics of nasal phonology: Theorems and data. In: Huffman, M. K., & Krakow, R. A. (eds.), *Nasals, Nasalization, and the Velum*, 5, 225–249. San Diego: Academic Press.

- Oliveira, C., Martins, P., Silva, S., & Teixeira, A. 2012. An MRI study of the oral articulation of European Portuguese nasal vowels. In: *Proceedings of INTERSPEECH 2012*.
- Orlikoff, R. F. 1991. Assessment of the dynamics of vocal fold contact from the electroglottogram: Data from normal male subjects. *Journal of Speech and Hearing Research*, 34, 1066–1072. DOI: <https://doi.org/10.1044/jshr.3405.1066>
- R Core Team. 2016. *R: A Language and Environment for Statistical Computing*. Vienna, Austria. Computer software program available from: <http://www.R-project.org>.
- Riordan, C. J. 1980. Larynx height during English stop consonants. *Journal of Phonetics*, 8, 353–360.
- Sampson, R. 1999. *Nasal vowel evolution in Romance*. Oxford: Oxford University Press.
- Shosted, R. 2015. Nasal vowels are not [+nasal] oral vowels. In: Smith, J., & Ihsane, T. (eds.), *Proceedings of the 42nd Linguistic Symposium on Romance Languages (2012)*. Amsterdam: John Benjamins. DOI: <https://doi.org/10.1075/rllt.7.05sho>
- Shosted, R., Carignan, C., & Rong, P. 2012. Managing the distinctiveness of phonemic nasal vowels: Articulatory evidence from Hindi. *Journal of the Acoustical Society of America*, 131(1), 455–465. DOI: <https://doi.org/10.1121/1.3665998>
- Shosted, R., Meneses, F., Pozzani, D., Wong, N., Hermes, Z., & Loucks, T. 2015. Articulatory characterization of oral/nasal vowel contrast in Brazilian Portuguese (São Paulo state): Evidence from EMA. In: *5th EICEFALA – International Meeting on Speech Sciences*.
- Solé, M.-J. 2007. Controlled and mechanical properties in speech: A review of the literature. In: Solé, M.-J., Beddor, P., & Ohala, M. (eds.), *Experimental Approaches to Phonology*, 302–321. Oxford: Oxford University Press.
- Stevens, K. N. 2000. *Acoustic phonetics*. Cambridge, MA: MIT Press.
- Straka, G. 1965. *Album phonétique*. Québec: L'Université de Laval.
- Styler, W. 2015. *On the acoustical and perceptual features of vowel nasality* (Unpublished doctoral dissertation). University of Colorado.
- Styler, W. 2017. On the acoustical features of vowel nasality in English and French. *Journal of the Acoustical Society of America*, 142(4), 2469–2482. DOI: <https://doi.org/10.1121/1.5008854>
- Teixeira, A., Martins, P., Oliveira, C., Ferreira, C., Silva, A., & Shosted, R. 2012. Realtime MRI for Portuguese. In: *Computational Processing of the Portuguese Language*, 306–317. Springer. DOI: [https://doi.org/10.1007/978-3-642-28885-2\\_35](https://doi.org/10.1007/978-3-642-28885-2_35)
- The Mathworks Inc. 2015. *Matlab 2015a*. Natick, Massachusetts, United States.
- Turner, R. L. 1921. Gujarati phonology. *Journal of the Royal Asiatic Society*, 3–4. DOI: <https://doi.org/10.1017/S0035869X00149287>
- Westbury, J. R. 1979. *Aspects of the temporal control of voicing in consonant clusters in English* (Unpublished doctoral dissertation). University of Texas, Austin.
- Wright, J. T. 1975. Effects of vowel nasalization on the perception of vowel height. In: Ferguson, C. A., Hyman, L. M., & Ohala, J. J. (eds.), *Nasálfest: Papers from a Symposium on Nasals and Nasalization*, 373–388. Palo Alto, CA: Stanford University Language Universals Project.
- Wright, J. T. 1986. The behavior of nasalized vowels in perceptual vowel space. In: Ohala, J. J., & Jaeger, J. J. (eds.), *Experimental Phonology*, 45–67. New York: Academic Press.
- Yu, A. C. L., & Lee, H. 2014. The stability of perceptual compensation for coarticulation within and across individuals: A cross-validation study. *Journal of the Acoustical Society of America*, 136(1), 382–388. DOI: <https://doi.org/10.1121/1.4883380>
- Zellou, G. 2017. Individual differences in the production of nasal coarticulation and perceptual compensation. *Journal of Phonetics*, 61, 13–29. DOI: <https://doi.org/10.1016/j.wocn.2016.12.002>

- Zemlin, W. R. 1998. *Speech and hearing science: Anatomy and physiology (4th ed.)*. Boston: Allyn and Bacon.
- Zerling, J. P. 1984. Phénomènes de nasalité et de nasalization vocaliques: Étude cinéradiographique pour deux locuteurs. *Travaux de l'Institut de Phonétique de Strasbourg*, 16, 241–266.

**How to cite this article:** Carignan, C. 2018 Using naïve listener imitations of native speaker productions to investigate mechanisms of listener-based sound change. *Laboratory Phonology: Journal of the Association for Laboratory Phonology* 9(1): 18, pp. 1–31, DOI: <https://doi.org/10.5334/labphon.136>

**Submitted:** 19 January 2018    **Accepted:** 09 August 2018    **Published:** 12 December 2018

**Copyright:** © 2018 The Author(s). This is an open-access article distributed under the terms of the Creative Commons Attribution 4.0 International License (CC-BY 4.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited. See <http://creativecommons.org/licenses/by/4.0/>.

]u[

*Laboratory Phonology: Journal of the Association for Laboratory Phonology* is a peer-reviewed open access journal published by Ubiquity Press.

**OPEN ACCESS** 