Title

On the acoustical features of vowel nasality in English and French

Journal

The Journal of the Acoustical Society of America, 142(4)

ISSN

0001-4966

Author

Styler, Will

Publication Date

2017-10-01

DOI

10.1121/1.5008854

On the acoustical features of vowel nasality in English and French

Will Styler[a)]

Department of Linguistics, University of Michigan, 611 Tappan Street, Ann Arbor, Michigan 48109, USA

(Received 6 February 2017; revised 23 August 2017; accepted 9 October 2017; published online 30 October 2017)

Although much is known about the linguistic function of vowel nasality, whether contrastive (as in French) or coarticulatory (as in English), and much effort has gone into identifying potential correlates for the phenomenon, this study examines these proposed features to find the optimal acoustic feature(s) for nasality measurement. To this end, a corpus of 4778 oral and nasal vowels in English and French was collected, and data for 22 features were extracted. A series of linear mixed-effects regressions highlighted three promising features with large oral-to-nasal feature differences and strong effects relative to normal oral vowel variability: A1-P0, F1's bandwidth, and spectral tilt. However, these three features, particularly A1-P0, showed considerable variation in baseline and range across speakers and vowels within each language. Moreover, although the features were consistent in direction across both languages, French speakers' productions showed markedly stronger effects, and showed evidence of spectral tilt beyond the nasal norm being used to enhance the oral-nasal contrast. These findings strongly suggest that the acoustic nature of vowel nasality is both language- and speaker-specific, and that, like vowel formants, nasality measurements require speaker normalization for across-speaker comparison, and that these acoustic properties should not be taken as constant across different languages. *© 2017 Acoustical Society of America.*

https://doi.org/10.1121/1.5008854

[CYE] Pages: 2469–2482

I. INTRODUCTION

Vowel nasality is a phenomenon of great importance in languages of the world, used both as an explicit carrier of phonological contrast in some languages, and as perceptual information for flanking nasal vowels. In French, as well as in Hindi, Lakhota, Navajo, and many other languages (constituting nearly 30% of the WALS dataset[1]), vowel nasality is phonemic, and a word's meaning can change depending whether the velum is raised or lowered during a given vowel's production. Thus, in Parisian French, *beau* [bo] means "beautiful" and *bon* [bõ] means "good." In these languages, the production and perception of nasality is doing obvious linguistic work, and is of crucial importance to the communicative act.

However, in languages where vowel nasality is not phonemically contrastive, as in English, vowel nasality is neither absent nor irrelevant, simply phonologically different. When speakers of any language produce nasal sounds in the vicinity of vowels, they *coarticulate*, that is, the nasal gesture overlaps the oral vowel gesture. Far from being simple asynchrony in articulatory timing, the resulting vowel nasality is still meaningful to listeners. Listeners use coarticulatory vowel nasality as a supplementary cue for oncoming nasals,[2,3] and there is strong evidence that nasality provides a cue for words with many phonological neighbors[4] and for easing difficult contrasts in non-words.[5] Even where nasality does not carry contrastive meaning, it is still attended to, and still useful to the listener as a disambiguating cue during speech perception. Given this importance, it is of descriptive interest and theoretical value (e.g., to theories of speech perception) that we are able to both understand and accurately measure the presence and degree of vowel nasality in natural human speech.

Although the literature on vowel nasality describes a wealth of acoustical correlates for nasality, no comprehensive effort has been made to directly compare the reliability and salience of these features, both for listeners in speech perception and for researchers when measuring nasality acoustically, nor has their reliability across different vowels and speakers been examined. This study, conducted as part of a larger experiment[6] seeking the perceptual cues to vowel nasality, takes this step by addressing fundamental questions about the acoustics of vowel nasality. Which acoustic features are the most useful indicators of vowel nasality? Are the measurements useful for comparison of nasality across speakers and across vowels? Do speakers of different languages in which nasality differs in phonological status use the same acoustical features when producing vowel nasality?

To provide a comprehensive look at the proposed features of nasality and address this gap in the literature, we examine 22 acoustical features here, some taken directly from the literature, some inferred from previous literature, and some newly proposed, in a corpus of oral, phonemically nasal, and coarticulatorily nasalized vowels in French and English. Using a series of statistical models, we identified those features which show both a statistically significant oral-to-nasal change, and evaluate the size of their effects, to determine which features are promising for measurement or perception.

We then examine the oral-to-nasal shifts in each feature across both the speakers and the vowels in the dataset. This

---

[a)]Electronic mail: will@savethevowels.org

will allow direct examination of the across-speaker and across-vowel variation in the dataset, to address the comparability of each feature from vowel to vowel and speaker to speaker, and to identify those features which are most promising for the measurement of nasality. Then, finally, we will examine the data on the whole and directly compare the results from English and French, to investigate whether nasality shows similar acoustic patterns in two different phonological forms: coarticulatory nasality (studied here in English) and contrastive nasality (here represented by French).

II. BACKGROUND AND FEATURE SELECTION

The literature on vowel nasality has generally discussed the acoustical consequences of nasality in terms of four types of acoustical change: the introduction of nasal resonances ("nasal poles"), the interference of these resonances with the oral resonances ("nasal zeros"), changes to the vowels' overall formant structures, and changes to the overall spectral envelope of the vowels. Spectral slices from two vowels, oral and nasal, are illustrated in Fig. 1, with many of the below features highlighted.

A. Nasal poles

The coupling of the complex nasal cavity with the oral cavity will necessarily add new resonances to the speech signal. These nasal resonances will result in regions of the spectrum where the harmonics are enhanced relative to the surroundings, resulting in "nasal peaks" or "nasal poles" or "nasal formants." These nasal poles will often be found in particular spectral regions (e.g., "P0" will tend to be found around 250 Hz for most speakers), and will affect the signal *in addition to* the existing oral resonances (e.g., formants). With this information in mind, we can discuss some of the most commonly referenced spectral poles attributed to nasality in the literature.

P0 is a low-frequency nasal pole, described in Ref. 7 as occurring "between 250 and 450 Hz" (p. 2360) with an amplitude increase between 3 and 5.5 dB, usually corresponding to the first or second harmonic (H1 or H2), although speakers with exceptionally short vocal tracts may have a higher P0. Chen attributes this peak directly to the resonant properties of the sphenoid and maxillary sinuses. This resonance is, in this author's experience, quite visually prominent when present, but in high vowels, where F1 strays into the 250–450 Hz range, or when the speaker's fundamental frequency is greater than 250 Hz, P0 cannot be readily identified nor measured. To avoid working with raw amplitude measures of harmonics (which are sensitive to recording conditions and word-by-word variation in speech volume), P0 is generally used in a relative measure as "A1-P0," where A1 is the amplitude of the highest harmonic in F1 and P0 (defined as the highest of H1 or H2) is subtracted from it. A1 is used as the second point in this relative measure because expect nasality to lower A1's amplitude (see Sec. II B) alongside the rise in P0, and thus, this relative measure combines two different acoustical elements of nasality in one number. Chen does offer an equation which attempts to "correct" the A1-P0 measure using the relative position and bandwidths of nearby formants when they enter P0's immediate vicinity, but in cases where F1 and P0 overlap directly, no information is directly recoverable. This "formant compensated" measure will be tested and referred to here as "A1-P0 (comp.)."

For cases of direct formant overlap, Ref. 8 suggests comparing A1 to P1, a second pole in the frequency range from 790 to 1100 Hz, with an average of 950 Hz. Although A1-P1 is usually considered a good alternative measure for high nasal vowels (where A1-P0 is not measurable), P1 is vulnerable to interference both from F1 and F2. A correction function based on the bandwidths and frequencies of nearby formants is again offered, but it too is only helpful if the overlap is not substantial enough to prevent the proper choice of P1. The interference from two separate formants and variation in its location mean that this peak can be exceptionally difficult to find, even by hand, and that consistent measurement of P1 can be quite difficult, more so than P0.

In addition, there exists a higher nasal pole, called P2, discussed in Ref. 9. Schwartz proposes that this peak is around 1250 Hz. No bandwidth nor amplitude estimates are offered, and again, no deterministic approach is offered for measurement, and the guideline given is only to find the highest harmonic in the 1250 Hz region in nasal vowels.

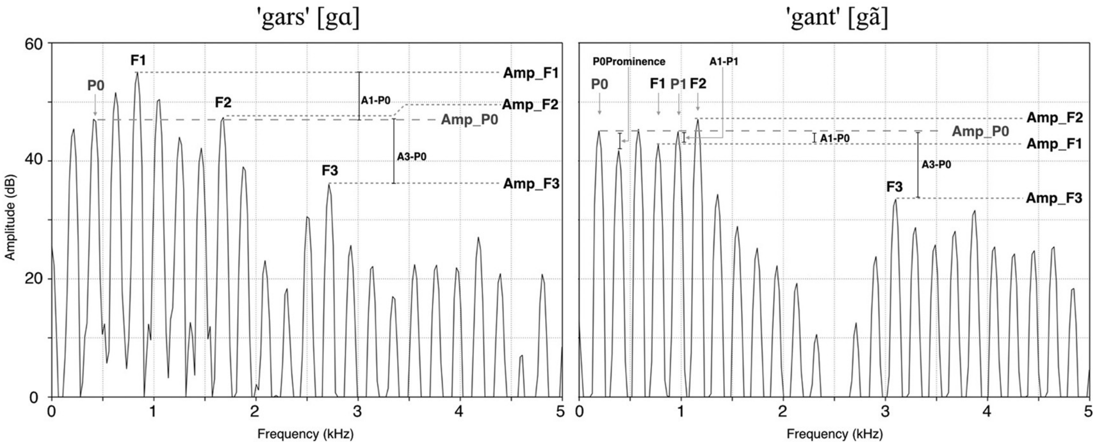

FIG. 1. Two spectra illustrating several nasal features at the 2/3 point of "gars" and "gant" for speaker fr7. Note that F1 overlaps P1's expected region in "gars."

The presence of a nasal pole does not necessarily entail that the frequency range in question will be higher in amplitude in nasal vowels. The overall reduction in spectral energy in nasal vowels, or the presence of a nearby zero, may reduce spectral energy more powerfully than the pole can add it back. In these cases, the harmonics associated with a nasal formant may have dropped in amplitude relative to the oral vowel, but have dropped less than the nearby frequencies.

To capture this phenomenon, we will also examine two features not previously described in the literature, P0 and P1 prominence. Defined as the local prominence of the nasal peak relative to the surrounding harmonics, that is, the amplitude of P0 minus the mean amplitude of the two immediately adjacent harmonics, this measure will hopefully capture the strong impression of the prominent "nasal peak" in spectra, even in those cases where the P0 harmonic is actually reduced in amplitude relative to oral vowels or where overall spectral energy is increased.

B. Nasal zeros

Alongside the poles created and described above, connecting the nasopharyngeal tract and nasal cavity can also result in reductions of amplitude in certain regions of the frequency spectrum, where the coupled resonances of the oral and nasal cavities result in destructive interference, creating what we refer to as "nasal zeros" or "nasal antiformants."

The strongest zero associated with nasality is in the region of the first vowel formant, and manifests as a sharp reduction of the amplitude of F1 relative to the surrounding spectral region. Damping in the region of the first formant is discussed extensively in the literature (as in Refs. 9–14, among many others). This drop in A1 also serves to enhance the pole measures discussed above (A1-P0, P1, or P2), as we are comparing a pole region which is meant to rise (or at least fall slowly) in nasal vowels to a zero region, which should lose power.

Alongside a drop in A1, a reduction in A2 and A3 (the amplitudes of the highest harmonics under F2 and F3) relative to the surrounding signal has also been proposed as a useful feature for the description of nasality (Refs. 9, 14, among others), and will be evaluated, again not as formant-specific effects, but as measures of spectral damping in the general regions of F2 and F3.

C. Formant changes

We have already discussed changes to vowel formant amplitudes due to nasal poles and zeros, but there is evidence that nasality modifies their frequency and bandwidth, as well.

First, we might expect differing formant structures in nasal vowels due to shifts in the oral articulation of nasal vowels. Some phonetic centralization of contrastive nasal vowels is far from unusual in language, and in many languages, phonologically allowable nasal vowels are a subset of the possible oral vowels in the language. But even when a nasal vowel is nominally in the same position (e.g., "high, front, unrounded"), studies report shifts in oral articulation relative to the oral vowel. References 11 and 15–17 all describe formant shifts in nasal vowels in French, Shosted and Carignan mention a similar effect in Hindi,[18] and both Refs. 13 and 19 describe formant shifts even in English, where nasality is not contrastive.

However, even if oral articulations did not shift in nasal vowels, the nasal poles and zeros could quite easily cause formant shifts, as changes in acoustic power near an existing oral formant will necessarily change its perceived width and central frequency. This alone could explain the sorts of formant shifts discussed in Refs. 14, 20, and 21 (among others). In addition, these nearby poles, influence of overlapping zeros, and heat loss due to the increased surface area in the nasal cavity can all lead to broadening of the formants (specifically F1), which has been mentioned repeatedly as consequence of vowel nasalization (in Refs. 13, 22–24, among others).

Thus, in this study, we will investigate the frequencies and bandwidths of the first three formants as potential features of nasality.

D. Broader spectral changes

The previously described poles, zeros, and formant shifts do not affect the vowel uniformly throughout the spectrum. The majority of nasal poles are low in the frequency range, while the zeros affect a broad swath of the spectrum, damping F1, F2, and F3. This uneven distribution of acoustical change necessarily alters the spectral tilt of the vowel, that is, the rate at which higher harmonics in a spectrum "fall off" in amplitude.

Although spectral tilt is primarily associated with voice quality in the phonetic literature, the link with nasality has been previously discussed. Reference 25 argues that spectral tilt, particularly the H1-H2 measure, is strongly affected by nasality and argues that low frequency spectral tilt is a poor feature for measuring voice quality as a result. In addition, Garellek *et al.*[26] have recently found a strong link between breathy voicing and nasality in three Yi languages. Although these voicing type changes are not a direct acoustical consequence of vowel nasality, such a secondary change could still be useful in identifying and measuring nasal vowels, and the data collected here can be used to corroborate this claim.

Thus, we will consider spectral tilt as a potential measure of nasality, using three features: H1-H2, spectral center of gravity, and a newly proposed feature for nasality-related spectral tilt, A3-P0. Although, for many speakers, this will correspond exactly to the comparisons of A3 and H1 occasionally used for spectral tilt in the voice quality literature (cf. Refs. 27 and 28, among others), this feature is designed to directly contrast the loss in amplitude of higher frequencies (captured by the amplitude of F3) with P0, the nasal pole.

E. Final feature set

Given these expected acoustical consequences of nasality from the literature and elsewhere, we can finalize the list of features for testing. Listed in Table I are the features to be evaluated, alongside their expected direction of change from oral to nasal vowels, and, where relevant, their provenance in the literature. Some additional features were

TABLE I. Features of nasality for evaluation, grouped by captured phenomenon, with expected direction of oral-to-nasal change and provenance in the literature.

Feature (change)	Description	Provenance
Zero features		
Amp_F1 (↓)	Amplitude of the 1st Formant ("A1")	(Ref. 11)
Amp_F2 (↓)	Amplitude of the 2nd Formant ("A2")	(Ref. 11)
Amp_F3 (↓)	Amplitude of the 3rd Formant ("A3")	(Ref. 11)
Pole features		
Amp_P0 (↑)	Amplitude of the "P0" nasal peak	(Ref. 7)
A1-P0 (↓)	Amp_F1-Amp_P0	(Ref. 7)
A1-P0 Comp. (↓)	A1-P0 using Chen's correction function	(Ref. 7)
P0Prominence (↑)	Prominence of P0 vs local harmonics	Not attested in literature
Amp_P1 (↑)	Amplitude of the "P1" nasal peak	(Ref. 8)
A1-P1 (↓)	Amp_F1-Amp_P1	(Ref. 8)
A1-P1 Comp. (↓)	A1-P1 using Chen's correction function	(Ref. 8)
P1Prominence (↑)	Prominence of P1 vs local harmonics	Not attested in literature
Amp_P2 (↑)	Amplitude of the "P2" nasal peak	(Ref. 9)
A1-P2 (↓)	Amp_F1-Amp_P2	Analogy from Ref. 7
Formant features		
Freq_F1	Frequency of the first formant (F1)	(Ref. 11)
Freq_F2	Frequency of the second formant (F2)	(Ref. 11)
Freq_F3	Frequency of the third formant (F3)	(Ref. 11)
Width_F1 (↑)	Bandwidth of the first formant	(Ref. 22)
Width_F2 (↑)	Bandwidth of the second formant	(Ref. 22)
Width_F3 (↑)	Bandwidth of the third formant	(Ref. 22)
Spectral features		
H1-H2 (↑)	Amp. of H1 minus Amp. of H2	(Ref. 25)
SpectralCOG (↓)	Spectral center of gravity of the vowel	Not attested for nasality
A3-P0 (↓)	F3's Amplitude minus P0's amplitude	Not attested for nasality

tested in Ref. 6 but, when both poorly performing and without precedent in the literature, were not reported here.

A subset of these features are shown graphically in Fig. 1, which shows two spectra, each taken at 2/3 of the vowel duration in the words "gars" and "gant" for French speaker fr7. Note the large decrease in A1-P0, the increase in spectral tilt (as A3-P0 and overall), the increased P0Prominence, and the strong reduction and widening of F1, F2, and F3. Also, note that although F1 is in the region where P1 should be measured in the oral vowel, we can see the P1 peak in the nasal vowel, and A1-P1 can be calculated.

III. METHODS

To evaluate these potential acoustical features of nasality, a sizable corpus of oral and nasal vowels was collected for both English and French. The word lists for these languages were designed to minimize differences in phonological environment.

A. Data collection—English

TABLE II. The English word list.

	CVC	CVN	NVC	NVN
/i/	deed	dean	need	neen
/i/	beeb	beam	meep	meme
/i/	bead	bean	mead	mean
/i/	deeb	deem	neeb	neem
/ɪ/	did	din	nit	ninny
/ɪ/	bib	bim	mib	mim
/ɪ/	bid	bin	mid	mint
/ɪ/	dib	dim	nib	nimble
/eɪ/	dade	deign	neighed	inane
/eɪ/	babe	bame	maybe	maim
/eɪ/	bayed	bane	maid	main
/eɪ/	dabe	dame	nape	name
/ɛ/	dead	den	ned	nen
/ɛ/	beb	bem	meb	memories
/ɛ/	bed	bent	meds	men
/ɛ/	deb	dems	neb	nem
/æ/	dad	dan	nad	nancy
/æ/	babble	bam	mab	ma'am
/æ/	bad	ban	mad	man
/æ/	dab	dam	nab	namble
/ɑ/	dodd	dawn	nod	non
/ɑ/	bob	bomb	mob	mom
/ɑ/	bod	bonfire	mod	monster
/ɑ/	dobson	dom	knob	nom
/aɪ/	died	dine	snide	nine
/aɪ/	imbibe	bime	mibe	mime
/aɪ/	bide	bind	mide	mine
/aɪ/	dibe	dime	nibe	nime
/ʌ/	dud	dunce	nudge	none
/ʌ/	bubba	bum	mub	mum
/ʌ/	bud	bun	mud	month
/ʌ/	dub	dumb	nub	numb
/oʊ/	doughed	don't	node	known
/oʊ/	bobe	bome	mobe	moam
/oʊ/	bowed	bone	mode	moan
/oʊ/	dobe	dome	noble	gnome
/u/	dude	dune	nude	noon
/u/	boob	boom	moob	moom
/u/	booed	boon	mood	moon
/u/	doobie	doom	noob	noom

For English, a series of quadruplets was solicited for each of the vowels /i, ɪ, eɪ, ɛ, æ, ɑ, aɪ, ʌ, oʊ, u/, consisting of words in the four coarticulatory structures examined in this work (CVC, CVN, NVC, NVN). These quadruplets each contained the same vowel paired with alveolar or bilabial onsets and codas (reducing the influence and interference of consonantal place cues on the vowel). Due to its restricted distribution, /ʊ/ was not recorded, and /ɔ/ and /ɑ/ were not distinguished due to their frequent merging in the local population. Where a necessary word for the tetrad did not exist in English, nonsense words like "neeb" or "mab" were recorded instead. As Ref. 5 showed little difference in degree and nature of nasality between non-words and real words, these need not be excluded from the analysis. The final word list is shown in Table II.

Speakers were recruited from the University of Colorado Department of Linguistics undergraduate subject pool. Twelve speakers were recorded, eleven self-identified as female and one as male, all native speakers of American English between the ages of 18 and 21. All English recordings took place in the University of Colorado Phonetics Lab, inside a sound-attenuated booth, using an Earthworks M30 microphone and Apogee Mini-Me Firewire Analog-to-Digital conversion box capturing audio at a 44 100 Hz sampling rate.

The words were isolated using the Penn Phonetics Lab Forced Aligner[29] and vowel boundaries were hand-confirmed by the author. Obstruent-vowel boundaries in (CVX and XVC) were identified using the waveform, looking for the first/final complete cycle at vowel-like amplitude following the onset/offset of periodicity. VN and NV boundaries were identified using both the waveform and spectrogram, identifying the nasal consonants using the change in waveform shape and amplitude, as well as the loss of amplitude in higher frequencies during nasal consonants.

The word list was recorded twice per speaker, presented in a single randomized ordering across all speakers in the carrier sentence "The word is X." This resulted in 160 tokens per read-through, with 2 read-throughs per speaker, yielding (in perfect circumstances) 320 tokens from each speaker, for a total of 3840 possible tokens. After discarding mispronunciations and other technical errors, we were left with 3823 tokens for analysis (99.6% of tokens collected).

B. Data collection—French

Although French has both coarticulatory and contrastive nasalization, the bulk of the literature and information on nasal acoustics in French which forms the foundation of this work examined phonemic nasality, and our choice to examine French and English is, in part, based on the desire to examine both coarticulatory and contrastive nasality. As such, and to limit the scope of the study and data, only contrastive nasality in French is studied here. As such, for French, 60 monosyllabic words in CV(C)/Cṽ(C) minimal pairs were collected, consisting of 10 oral-nasal pairs for each of the vowels /ɔ, a, ɛ/. The final word list is shown in Table III.

Six speakers were recorded in the University of Colorado Phonetics Lab (using the same equipment described above), and two additional French speakers (7 and 8) were recorded in the University of Pennsylvania Phonetics lab, using comparable equipment. All speakers were between the ages of 23 and 45, and were born and raised in Northern France. Of these, three self-identified as male, and five as female, and all were at least proficient in English.

The words were isolated using the SPLAligner French Forced Alignment tool from Peter Milne (described in Ref. 30), and vowel boundaries were again hand-verified, using the same criteria for English for obstruent boundaries. For fricative boundaries, the onset of modal voicing of vocalic amplitude was used, and for sonorant boundaries (/ʀ/ and /l/), both the spectrogram and waveform were used, using changes waveform amplitude and formant structure to isolate the vocalic portions.

The French words were recorded in the carrier sentence "Dites [word] s'il vous plaît" ("Say [word] please"). Although Ref. 31 notes that fricatives can lead to spontaneous nasalization even in oral vowels, potentially giving rise to nasality in oral tokens prior to "s'il," we should note that the carrier phrase is uniform across tokens, and effects of any spontaneous nasalization would work against the hypotheses tested here (that these features are linked to nasality), as it would reduce the change in feature from oral to nasal vowels. Tokens were presented for recording in a single, randomized ordering to all speakers, with two repetitions of each word. This resulted in 120 words per speaker being recorded. With mispronunciations and errors discarded, this resulted in 955 usable tokens (99.5% of recorded tokens).

TABLE III. The French word list.

	Nasal word	IPA	Gloss	Oral pair	IPA	Gloss
ɛ̃	train	tʀɛ̃	train	trait	tʀɛ	trait
ɛ̃	lin	lɛ̃	linen	lait	lɛ	milk
ɛ̃	rein	ʀɛ̃	kidney	rai	ʀɛ	ray (of light)
ɛ̃	crin	kʀɛ̃	hair	craie	kʀɛ	chalk
ɛ̃	prince	pʀɛ̃s	prince	presse	pʀɛs	press
ɛ̃	plein	plɛ̃	full	plaie	plɛ	wound
ɛ̃	inde	ɛ̃d	India	aide	ɛd	help
ɛ̃	gains	gɛ̃	profits	gay	gɛ	gay man
ɛ̃	fin	fɛ̃	end	fait	fɛ	fact
ɛ̃	bain	bɛ̃	bath	baie	bɛ	bay
ɑ̃	pan	pɑ̃	section	pas	pɑ	pace
ɑ̃	temps	tɑ̃	weather	tas	tɑ	heap
ɑ̃	planque	plɑ̃k	hideout	plaque	plɑk	sheet
ɑ̃	plan	plɑ̃	plan	plat	plɑ	dish
ɑ̃	gland	glɑ̃	acorn	glas	glɑ	knell
ɑ̃	gant	gɑ̃	glove	gars	gɑ	guy
ɑ̃	chance	ʃɑ̃s	luck	chasse	ʃɑs	chase
ɑ̃	camp	kɑ̃	camp	cas	kɑ	case
ɑ̃	grande	gʀɑ̃d	big one	grade	gʀɑd	rank
ɑ̃	ban	bɑ̃	cheer	bas	bɑ	quietly
ɔ̃	ronce	ʀɔ̃s	bramble branch	rosse	ʀɔs	nag
ɔ̃	rhombe	ʀɔ̃b	musical instrument	robe	ʀɔb	dress
ɔ̃	ponte	pɔ̃t	clutch	pote	pɔt	buddy
ɔ̃	pompe	pɔ̃p	pump	pop	pɔp	pop music
ɔ̃	onde	ɔ̃d	wave	ode	ɔd	ode
ɔ̃	once	ɔ̃s	ounce	os	ɔs	bone
ɔ̃	honte	ɔ̃t	shame	hot	ɔt	hot
ɔ̃	conque	kɔ̃k	conch shell	coq	kɔk	rooster
ɔ̃	comte	kɔ̃t	account	cote	kɔt	quotation
ɔ̃	bombes	bɔ̃b	bombs	bob	bɔb	sun-hat

C. Analysis

1. Feature extraction

All features were measured automatically by script in Praat.[32] Each measurement was taken at two points per vowel, at 1/3 and 2/3 of the vowel's duration, such that both carryover (NVC, NVN) and anticipatory (CVN, NVN) nasality could be adequately captured. Once a spectrum was generated, formant frequency and amplitude (A1, A2, and A3) measures were determined by finding the amplitude and frequency of the highest-amplitude harmonic under or near the formant center given by the PRAAT Built-in LPC analysis (five formants in 5 kHz), and bandwidth was taken directly from the LPC estimate. The amplitude and frequency of individual harmonics (e.g., H1, H2, P0, P1) were found by identifying the strongest spectral peak in the harmonic's expected region based on the F0 track. For P0, the higher of H1 or H2 was used, as suggested in Chen's original description of the measure. For P1, the strongest harmonic was found in the 850–1050 Hz range (again, using the range suggested by Chen[8]). For P2, the strongest harmonic between 1150 and 1350 Hz was used, using the range outlined in Ref. 9, the primary description of the feature. Compensation for A1-P0 and A1-P1 was applied using the formulae in Ref. 7 using bandwidths and frequencies from the LPC analysis. Finally, P0Prominence and P1Prominence were calculated by taking the difference between the previously identified nasal peak and the mean of the surrounding harmonics.

The automatic measurement PRAAT script, available for inspection and use (see Ref. 33), included a series of steps and automatic, internal checks which were developed after extensive visual comparison of measured values with the generated spectra, to help avoid known failure modes of spectral analysis and reliably identify the same peaks as a human annotator would visually. To avoid the influence of nearby formant and pitch shifts, and to provide cleaner harmonics, each spectral slice was generated after iterating the temporally specified voicing cycle to meet the required window length. In addition to excluding measurements portions of the vowel where F0 could not be reliably tracked (often due to creaky voice), for this group of speakers in these data, only F0 values between 80 and 300 Hz were allowed, and P0 was constrained to harmonics under 500 Hz, to avoid choosing a phonetically implausible harmonic when H2 is high. To avoid formant tracking errors (e.g., F2 found as F1), F1 was checked to ensure that it occurred only in a reasonable space for our speakers (180–1000 Hz). When the identified F1 peak exceeded these thresholds, the script automatically re-measured the vowel using fewer formants, and excluded the vowel if this subsequent measurement also failed. To rule out cases of pitch doubling, measurements were automatically re-checked at an immediately adjacent point if the found H1 was greater than twice the vowel's average F0. Because individual harmonics were found by finding the highest peaks in the expected region, in order to ensure that we found H2 (rather than a local maximum), the H2 harmonic was checked to ensure that it was within half an F0 of F0*2. Then, finally, the depth of the valley between H1 and H2 was verified to be at least 5 dB, to ensure that the spectrum clearly differentiated harmonics, and to avoid capturing a "notch" or local medium. Finally, once each speaker's measurements were extracted, individual measurements were excluded from analysis which fell implausibly outside each participant's normal range for F1, F2, F3, and for F0 (with a threshold at 3 standard deviations from the participant's by-vowel mean for the measure).

Note that these same methods were applied for finding P0, P1, and P2 (and all other measures) in oral vowels as well, even though we do not expect to find actual nasal poles in oral vowels. By holding the measurement process constant regardless of expected nasality, that is, identifying harmonics as "P0," "P1," and "P2" even in oral vowels using the same "highest harmonic in the expected region" method, we are able to apply these measurements uniformly, and can directly compare values in vowels of any known (or unknown) nasality. Any measure(s) which cannot be applied to oral vowels (e.g., due to excessive noise in extraction) will prove ineffective in our final analysis, and will be eliminated from further consideration.

2. Oral and nasal

The goal of the present work, to identify features which change meaningfully and predictably with vowel nasality, is relatively straightforward, but it is complicated by the fact that 22 features must be examined. To make this analysis tractable and readable, we will simplify the process in several ways.

First, we will take a very basic approach to establishing "oral" vs "nasal." In the French data, "oral" (e.g., "fait" (/fɛ/) and "nasal"(e.g., "fin" (/fɛ̃/) correspond to the phonological categories. The link between articulatory nasality and phonological nasality has been shown repeatedly in both airflow and acoustics (cf. Ref. 34 among others). Thus, in French, for each feature, we will compare the measured values between the oral vowel ($\text{nasality} = 0$) and the nasal vowel ($\text{nasality} = 1$) items of each minimal pair, at two timepoints per word.

In English, vowel nasality is triggered by adjacent nasal consonants. Thus, CVC words can be assumed to be "oral" for our analysis. The "nasal" category is a bit more complex. Any vowel with surrounding nasal context will have *some* nasal influence. As shown using airflow studies (cf. Ref. 34), there is often some degree of nasal airflow *throughout the vowel* in CVN, NVC, and NVN contexts. Although NVN syllables will be most consistently nasal, nasality should consistently be found in CVN and NVC vowels, and never in CVC. It is also worth noting that the effect of any non-nasal points measured in these contexts will be conservative, favoring the null hypothesis that a feature has no link to nasality. This is borne out in tests which showed NVN-only data to show similar patterns of significance, but with slightly higher coefficients. Thus, although the categories will differ in nasality for some speakers and tokens, classifying CVN, NVC, *and* NVN vowels as "nasal" will triple the number of comparisons possible, and the overall effect will favor a careful interpretation of the results.

So, for English, we will compare the measured values for each feature between CVC-context vowels ($\text{nasality} = 0$) and CVN/NVC/NVN-context vowels ($\text{nasality} = 1$), again measuring two timepoints per vowel.

3. Statistical analysis

To identify features meaningfully associated with vowel nasality, we examine the change in each measure between oral and nasal vowels, henceforth referred to as ΔFeature. We will establish the significance of ΔFeature for each feature using a linear mixed effects regression (henceforth, "LMER"), as implemented in the lme4 package in R.[35]

To measure the statistical strength of each ΔFeature as outlined above, for each language and feature, an LMER was run, including timepoint as a fixed effect, random intercepts for word, and random slopes for speaker and vowel (as we would expect each to vary both in degree and amount of change). So, for example, to evaluate the relationship between the amplitude of A1 ("Amp_F1") and nasality in English, the R code below would be run (noting the difference in syntax between random and fixed effects, as well as the by-speaker and by-vowel random slopes for nasality),

```
Amp_F1.lmer = lmer(Amp_F1 ~ nasality
                   + Timepoint + (1|Word
                   + (1 + nasality|speaker)
                   + (1 + nasality|vowel),
                     data = eng))
```

This outputs coefficients for all fixed effects (one each for nasality and timepoint), as well as the variance absorbed by word, speaker, and vowel. In addition, for each fixed effect, the t statistic is examined to determine the significance of ΔFeature, taking $|t| > 2$ as an indicator of statistical significance (after Ref. 36). Although each model will generate a full set of coefficients and complete output, in the interest of space, this additional output will not be included or discussed *en masse* here.

It is worth noting that this analysis evaluates each feature independently, and treats each feature as a monolithic measure. Although many of the features evaluated are closely related to one-another (Amp_F1, Amp_P0, and A1-P0, for instance), the goal of the present work is not to attempt to fully explain the holistic acoustical change associated with nasal coupling, but instead, to identify individually measurable features which are correlated with vowel nasality. It is the case that "clusters" of related features will show relationships with nasality, but our goal here is to identify those which show the clearest ΔFeature.

Finally, a word on effect size. Even if there is a statistically significant ΔFeature between oral and nasal vowels, a given measure is not practically useful unless the ΔFeature is greater than the baseline variability in that feature (that is, the variation in non-nasal or CVC vowels). To evaluate the relative "strength" of each feature in nasal vowels relative to this variation and noise in known-oral vowels, we will be discussing the "CoefVsSD" value, that is, the model-derived coefficient for nasality (ΔFeature) divided by the raw standard deviation of the measure *in oral/CVC tokens only*. We use the raw, model-external standard deviation here as it represents a worst-case scenario for variability, where all of the (presumably tractable) by-speaker, by-vowel or by-token variation is included without control, yielding a more conservative measure than a model-internal version.

This process yields a single number, a sort of signal to noise ratio, which allows us to compare the various measures in terms of their practical utility for indicating nasality, regardless of the unit or language. A very low CoefVsSD indicates that the oral-to-nasal ΔFeature may easily be "lost" in the token-by-token variation expected for that feature, where a higher number indicates that the nasality-associated ΔFeature will be more often able to rise above the normal variation. By examining this number, we can get a rough sense of the utility of the feature for differentiating oral and nasal vowels. Although there is not a strict "cut off" below which a feature will not be considered useful or meaningful, this measure can help us to identify those features which are not only statistically significant, but whose oral-to-nasal effects are large enough relative to the normal variation to be of reliable use for nasality measurement or perception.

IV. RESULTS—ENGLISH

The results for English are presented in Table IV, ranked by $|t|$. Thirteen of the 22 features reached the $|t| > 2$ threshold for significance in the English dataset. For each feature, we present the model-derived coefficient for nasality from the model (ΔFeature from oral to nasal), the t value of the oral-to-nasal ΔFeature, as well as the means for the feature in both oral and nasal words, and the CoefvsSD measurement described above. Rather than discussing each individually, we will focus on clusters of related features, and reserve individual comment for the best performers.

Reassuringly, the A1-P0 cluster of features (A1-P0 Compensated, A1-P0, Amp_F1, Amp_P0, and P0Prominence) A1-P0 features will necessarily show a decrease with nasality, as expected from the literature. We also see in Table IV that Chen's compensation algorithm provides a small improvement over raw A1-P0, although the difference is minimal. In addition to confirming the statistical link of these known features, we can also see that the A1-P0 variants show a stronger signal relative to oral noise (as shown by CoefVsSD) than either F1

TABLE IV. English model output by feature, including the oral-to-nasal coefficient (ΔFeature), t-statistic for nasality, oral and nasal means, and effect size as measured by CoefvsSD.

Features	Nasality coef.	Nasality t	Nasal mean	Oral mean	CoefvsSD
A1-P0 (comp.)	−4.119	−7.185	1.133	5.099	0.571
A1-P0	−4.071	−6.998	−0.982	2.932	0.563
Amp_F1	−2.534	−6.753	43.962	46.445	0.376
Width_F1	96.075	6.083	262.532	171.093	0.478
A3-P0	−3.585	−4.984	−18.179	−14.965	0.347
P0Prominence	1.863	4.181	12.083	10.290	0.262
Amp_P0	1.539	3.799	44.944	43.513	0.249
A1-P2	−3.131	−3.132	14.422	17.050	0.265
Amp_F3	−2.061	−3.086	26.765	28.547	0.212
A1-P1 (Comp.)	−3.532	−2.884	16.137	19.398	0.325
SpectralCOG	−41.125	−2.143	717.497	758.323	0.120
A1-P1	−3.157	−2.128	12.801	15.653	0.252
Width_F3	97.163	2.068	589.278	498.014	0.137
H1-H2	1.128	1.550	4.341	3.217	n.s.
Freq_F1	31.931	1.433	629.036	601.869	n.s.
P1Prominence	0.768	1.067	2.043	1.458	n.s.
Amp_P2	0.577	0.536	29.540	29.395	n.s.
Freq_F2	−24.586	−0.520	1762.837	1770.088	n.s.
Amp_F2	−0.294	−0.478	35.259	35.768	n.s.
Amp_P1	0.626	0.401	31.161	30.792	n.s.
Width_F2	15.525	0.314	442.663	421.553	n.s.
Freq_F3	8.480	0.233	2,775.772	2,756.662	n.s.

or P0's amplitudes alone, pointing to the utility of the composite measure, despite the increase in complexity.

The bandwidth of F1 also showed a strong link to nasality, showing a nearly 100 Hz widening in nasal contexts (t = 6.08). Of note, bandwidth was not significant for F2, and was only barely so (t = 2.07) for F3, and even then, with a very small effect size relative to oral variation, pointing to a particular affinity between nasality and F1's bandwidth.

Spectral tilt proved strongly linked to nasality as well, with both SpectralCOG and A3-P0 showing significant ΔFeature, indicating that energy falls off more quickly in nasal than in oral vowels, although A3-P0 showed a much stronger effect relative to oral noise, perhaps owing to F3's significant drop in amplitude (−2 dB, t = −3.09). H1-H2, however, failed to reach significance, perhaps owing to the strong variation in H1 and H2's amplitude due to the changing correspondence of H1, H2 and P0 across speakers.

Two members of the A1-P1 Cluster (A1-P1 and A1-P1 Compensated) showed significant ΔFeature in nasal vowels. Of particular interest is the fact that A1-P1 reached significance in this dataset, containing both high *and* low vowels, where A1-P1 is generally used only for high vowels. However, neither P1's Prominence nor amplitude reached significance, indicating that this effect, too, may mostly be driven by the drop in A1.

A1-P2 also showed a significant ΔFeature. Again, though, this had a notably weaker effect than A1-P0 or even F1's amplitude alone in terms of signal-to-oral-noise, and again, P2's amplitude did not reach significance in this dataset, again suggesting that A1's amplitude drives this effect.

Finally, we also see that the frequency of F1, F2, and F3 *across all vowels* was not significantly correlated with nasality. This indicates that although individual vowels may show nasality-related changes in formant structure, perhaps related to differing oral articulations of nasal vowels, the velopharyngeal coupling associated with nasality *itself* does not affect formant frequencies in a consistent manner across vowels. Put differently, although individual vowels may differ, there are no formant frequency changes *to the entire vowel space* associated with nasality.

In addition to the significance of the oral-to-nasal ΔFeature, we should also consider the magnitude of the change. Although none of these measures show an oral-nasal change greater than one standard deviation for the oral vowel, we can see that, for instance, a 95+ Hz bandwidth change is far more unusual (in terms of the variability found in oral vowels) for F1 than for F3, and that A1-P0-related measures are particularly strong relative to oral token-by-token variation. By this metric, we see that although 13 of the measures reached significance, the strongest and most reliably measured spectral features were A1-P0 (and components), F1's Bandwidth, spectral tilt, and A1-P1. Note that this (extremely conservative) measure uses model-external standard deviations, and thus, the variability used for comparison represents a worst-case scenario.

It is also worth mentioning that all but two features found significant above still show significant effects in a simplified CVC/NVN comparison (results given in Table V), although the coefficients are universally higher (representing the greater and more consistent degree of nasality in NVNs). Those two features which lost significance (Width_F3 and A1-P1) were only marginally so in the main model, and their loss of significance can be attributed to the reduction in N (and thus, of power). But given this result, we can assume that the ΔFeature values reported above are fairly conservative, and ΔFeature in high-nasality contexts (NVNs) will likely be higher.

TABLE V. English CVC vs NVN model output by feature, including the oral-to-nasal coefficient (ΔFeature), t-statistic for nasality, oral and nasal means, and effect size as measured by CoefvsSD.

Features	Nasality coef.	Nasality t	NVN mean	CVC mean	CoefvsSD
A1-P0 (Comp.)	−5.424	−7.611	−0.162	5.099	0.752
A1-P0	−5.368	−7.415	−2.271	2.932	0.742
Amp_F1	−3.438	−7.232	42.904	46.445	0.510
Width_F1	128.826	6.848	296.747	171.093	0.642
P0Prominence	2.966	5.799	13.176	10.290	0.417
A3P0	−5.333	−5.458	−19.629	−14.965	0.516
Amp_P0	1.935	3.874	45.176	43.513	0.314
Amp_F3	−3.400	−3.693	25.546	28.547	0.349
A1-P1 (comp.)	−4.303	−2.639	15.420	19.398	0.396
SpectralCOG	−76.891	−2.571	684.064	758.323	0.223
A1-P2	−3.828	−2.399	14.039	17.050	0.325
Width_F3	131.070	1.987	615.589	498.014	n.s.
A1-P1	−3.839	−1.969	12.207	15.653	n.s.
P1Prominence	1.500	1.849	2.826	1.458	n.s.
Freq_F1	36.343	1.138	632.548	601.869	n.s.
Amp_F2	−0.605	−0.675	34.595	35.768	n.s.
H1-H2	0.675	0.647	3.651	3.217	n.s.
Freq_F2	−42.637	−0.628	1762.474	1770.088	n.s.
Freq_F3	15.159	0.280	2795.524	2756.662	n.s.
Amp_P2	0.373	0.218	28.865	29.395	n.s.
Amp_P1	0.401	0.195	30.697	30.792	n.s.
Width_F2	−0.675	−0.010	438.916	421.553	n.s.

V. RESULTS—FRENCH

The results of the French study are presented in Table VI, again ranked by |t| and using the same columns and data format. In French, only 10 of the 22 features reached the threshold for significance.

The strongest set of features for the French speakers all measure spectral tilt, whether directly or indirectly. A3-P0 and Spectral Center of Gravity both indicate changes in spectral tilt, in this case, a strong damping of higher frequencies. This is mirrored by changes in F3's amplitude and bandwidth (lower and wider, respectively, in nasal vowels), as well as by H1-H2. Thus, particularly considering the strength of these changes relative to noise in oral vowels (CoefvsSD), we see that spectral tilt is the strongest indicator of nasality in these French data.

As in English, the A1-P0 complex of measures shows a strong change between V and Ṽ words, with A1-P0 (in both forms), Amp_F1, and P0Prominence all reaching significance. And again, these ΔFeature values showed promisingly strong differences relative to the variability in oral vowels, and as in English, F1's Bandwidth showed both significant and strong ΔFeature in nasal vowels.

TABLE VI. French model output by feature, including the oral-to-nasal coefficient (ΔFeature), t-statistic for nasality, oral and nasal means, and effect size as measured by CoefvsSD.

Features	Nasality coef.	Nasality t	Nasal mean	Oral mean	CoefvsSD
A3-P0	−10.293	−6.600	−22.816	−12.537	1.161
Amp_F3	−11.293	−5.615	15.154	26.579	1.013
Width_F3	359.620	3.522	760.728	400.098	0.653
A1-P0 (comp.)	−5.632	−3.318	2.932	8.681	1.100
A1-P0	−5.665	−3.287	0.719	6.500	1.105
Amp_F1	−6.669	−3.251	38.690	45.616	0.666
SpectralCOG	−143.133	−2.636	636.496	779.721	0.545
Width_F1	157.304	2.572	365.041	205.784	0.432
H1-H2	3.135	2.484	0.170	−2.962	0.456
P0Prominence	2.673	2.362	5.645	2.997	0.424
P1Prominence	0.844	1.598	1.206	0.378	n.s.
Amp_F2	−7.475	−1.184	28.181	35.843	n.s.
Width_F2	228.899	0.999	605.643	374.930	n.s.
Amp_P0	−1.003	−0.979	37.970	39.116	n.s.
Freq_F2	−179.928	−0.727	1410.198	1588.074	n.s.
Amp_P2	−6.636	−0.690	21.486	28.237	n.s.
Amp_P1	−4.634	−0.689	28.478	33.322	n.s.
A1-P1 (comp.)	−1.828	−0.378	13.764	15.657	n.s.
A1-P1	−2.030	−0.368	10.211	12.294	n.s.
Freq_F3	54.771	0.363	2729.481	2676.358	n.s.
Freq_F1	−24.044	−0.295	613.365	637.208	n.s.
A1-P2	−0.040	−0.004	17.203	17.379	n.s.

Of note, although both A1-P1 (in all forms) and A1-P2 showed reasonable performance in English, neither is significant in French. This is not surprising in light of the datasets: The English data included large numbers of mid-high and high vowels, where these two measures are most useful, while the French dataset included only the mid and low vowels in which A1-P1 and A1-P2 have frequent interference from F1.

As in English, across-vowel formant shifts did not prove significant in these data, indicating again that although there may well be vowel-specific formant changes, there is no overall formant shift associated with nasality.

Finally, it is worth noting that CoefvsSD values are generally higher for French than for equivalent measures in English. Although some reduction in variability can be attributed to a reduction in the number of speakers and vowels, the sharp increase in signal-to-noise indicates greater difference in these features between oral and nasal vowels in French than in English.

VI. DISCUSSION

With these data in hand, we can answer our four main research questions. First, we will evaluate which of the analyzed features are most useful for measuring nasality. In this process, we will examine the across-vowel and across-speaker variability in the best candidates, to further gauge their reliability. Then, finally, we will directly compare the French and English data, and examine the extent to which the acoustic realization of vowel nasalization is similar (or different) across these two languages.

A. Which acoustical features are best for measuring nasality?

For a particular measure to be useful in the study of nasality, it needs to have two characteristics. First, the feature must show a statistically significant correlation with nasality, showing a meaningful ΔFeature between oral and nasal vowels. Second, this ΔFeature must be meaningfully large given the feature's normal variation in oral vowels.

Given these criteria, we can narrow the field from the initial 22 features to three relatively independent acoustical phenomena, represented by four specific acoustic features.

1. Poles and zeros (A1-P0 and A1-P1)

First, the A1-P0 cluster of features (namely, A1P0, Amp_F1, Amp_P0, and P0Prominence) is very strong in both English and French, indicating that the low-frequency pole-zero pair is robust and detectable. Given the findings of Refs. 7 and 8, as well as the many subsequent studies using A1-P0, this should come as no surprise. Of the various measures associated with this pole-zero pair, Chen's A1-P0 measure shows the largest ΔFeature and the strongest effect relative to noise in oral vowels. Indeed, when considering languages together, A1-P0 is the strongest performer, and we find its common use in the literature well supported, even in high vowels where it is often avoided. Note, though, that although A1-P0's performance is slightly increased beyond the raw A1-P0 value when Chen's formant compensation calculation is used in the 11-vowel English dataset, this advantage disappears when the number of vowel qualities is reduced to 3 in French. Thus, the author takes no strong position about the utility of this compensation calculation.

The A1-P1 cluster of features has a mixed showing in these data. A1-P1 (and its formant-compensated analog) showed a significant ΔFeature in English, but not in French, and even in English the effect was not strong, as ΔFeature was weak relative to the feature's variation in oral vowels. The cross-linguistic difference in A1-P1's effectiveness can be attributed, in part, to the lack of high vowels, where F1 and P1 do not interact, but ultimately, the feature is not strong in either language, as even in English, the signal-to-oral-noise is greater when we look at A1's amplitude alone. This, coupled with the relative difficulty of identifying a clear "P1" for some speakers' vowels, leaves this feature useful in some situations, but far from ideal.

Finally, A1-P2's utility as a feature for measuring or identifying nasality is not supported in these data. In French, no P2-related feature reached significance, and in English, although A1-P2 showed a significant effect, both the ΔFeature and the effect size relative to variation in oral vowels were weaker for A1-P2 than for Amp_F1 alone, suggesting that incorporating the P2 peak into the measure was actually counterproductive.

2. Formant changes (formant bandwidth)

In both languages, the bandwidth of F1 (and to a lesser extent F3) proved a strong acoustical feature of nasality, with nasal vowels showing a significant widening of the formants, far greater than the normal variability in oral vowels. Although some of the broadening likely results from the nasal zero's reduction of F1's amplitude, it also makes considerable acoustical sense, given the increased thermal and surface loss of energy with the increased volume when the nasal cavity is coupled into the system. F1's bandwidth, then, merits further use as a measurement of nasal coupling.

3. Overall spectral changes (spectral tilt)

In both languages, nasal vowels showed a strong increase in spectral tilt, with the harmonic structure in nasal vowels falling off more rapidly than in oral vowels. This effect was represented in both languages across several different measures (H1-H2, spectral center of gravity, reductions in higher frequency peaks, and the purpose-built A3-P0). In French, this change in spectral tilt (via A3-P0) showed the strongest effect of *any* of our 22 features. Although some of this can be attributed to an increase in the lower frequency nasal pole and the nasal zeros, the reduction in higher frequencies is more profound than can be explained by the pole-zero complex alone. This increased ΔFeature, particularly in French, points to the possibility that an increase in spectral tilt (beyond that caused by the pole-zero complex) has been recruited to enhance the contrastiveness of the oral-nasal distinction in French (and potentially in other languages as well).

However, despite the strength of its statistical relationship to nasality, caution must be used here. Spectral tilt is strongly affected by many non-nasal factors in speech, the most common being voicing type (cf. Refs. 37 and 38), stress (cf. Ref. 39), and vocal pathology (cf. Ref. 40). So, although it may be useful as a secondary feature of nasality for measurement, particularly in languages where it shows a stronger ΔFeature in nasal vowels, A3-P0 is affected by too many non-nasal factors to be useful as a specific measure of nasality. This further underscores the main point of Ref. 25, indicating that not only does nasality complicate the use of H1-H2 to measure voice quality, but that nasality may interact with spectral tilt "on the whole," interfering even with broad-spectrum voice quality measures like A3-H1.

B. Vowel variability

First, we must highlight an important distinction. It is well known that nasal vowels differ in their oral articulations from oral vowels (cf. Refs. 11 and 15–17), and these effects are specific *to each individual vowel*, within each language. That is, there are differences in a nasalized /a/ which are not directly caused by the acoustic coupling, and would not be found in a nasalized /ɛ/.

Rather than focusing on the differences between oral and nasal instances of *any particular vowel*, the present study aimed to find direct acoustic consequences of vowel nasality, that is, acoustic phenomena which are caused directly by the oral-nasal coupling, applicable to any vowel. Although we expect some by-vowel variation in spectral measures simply because of the differing formant structure of different vowels, for a feature to be of interest here, it should show a robust effect, in the same direction, across all vowels examined.

Figure 2 shows the oral and nasal mean, for each vowel in English and then French, across our four most promising acoustic features. For each vowel, we see the average for each feature in both oral and nasal vowels in our dataset. For instance, we see that /ɔ̃/ in French had a mean A1-P0 of ∼5 in oral vowels, and ∼0 in nasal vowels.

First, we can see that A1-P0 and F1's bandwidth were uniformly useful, with all vowels showing meaningful ΔFeature in the expected directions, albeit with different baselines and ranges. Spectral tilt (A3-P0) is also shown to be uniform across these three vowels, although again varying in range and baseline according to vowel quality. We also see that, mirroring the statistical models, the change in spectral tilt associated with nasality is a great deal stronger in the French dataset relative to the English data.

A1-P1 displays a great deal of across-vowel variability, ranging from acceptable results (in mid vowels) to showing no meaningful ΔFeature, and interestingly, performed exceptionally poorly in the high vowels, where it is generally regarded as a safer option than A1-P0. Although some of this variability in usefulness likely stems from difficulty identifying the "proper" P1 peak, given that P1 is measured

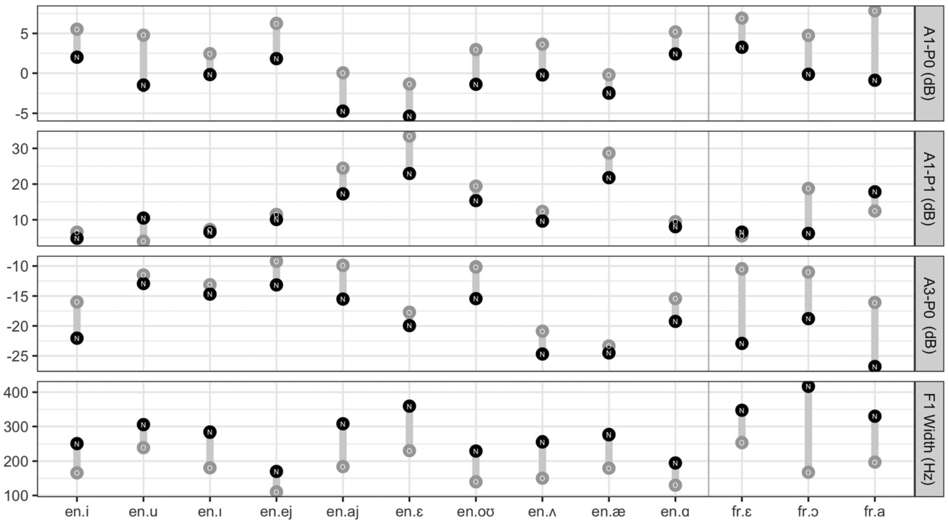

FIG. 2. Oral (gray) and nasal (black) means for each of four features across each of the vowels, in English (left, marked "en") and French (right, marked "fr"). Features scaled individually.

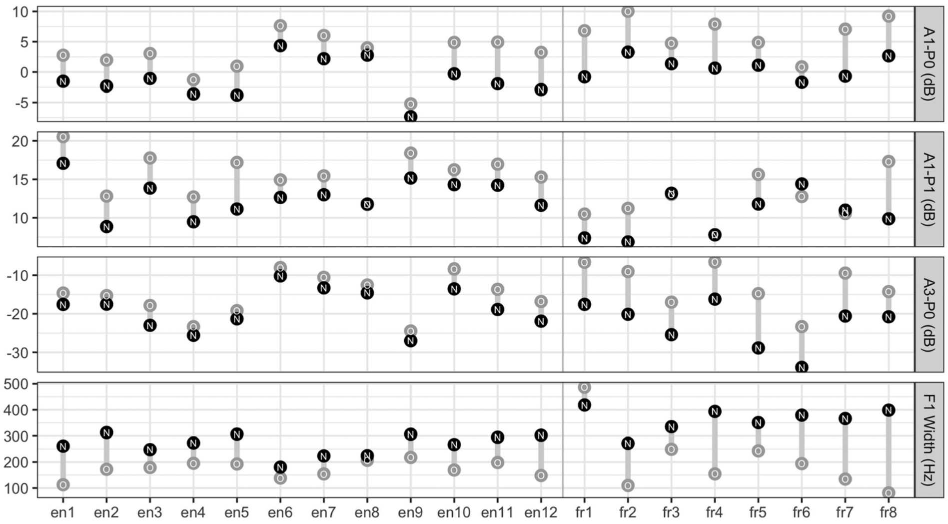

FIG. 3. Oral (gray) and nasal (black) means for each of four features across each of the speakers, in English (left, marked "en") and French (right, marked "fr"). Features scaled individually.

automatically by finding the highest peak in P1's given range (per Ref. 8), the extraction process and code used was identical to that for A1-P0, simply using a different frequency range, so the difference between these features cannot be explained entirely by a "bug" or an improper peak selection heuristic. As many who have worked with the measure can attest, A1-P1 can be exceedingly difficult to measure, even by hand, and some speakers and vowels appear to lack an identifiable P1, even in known-nasal contexts, so it is not surprising that P1 would show significant variability and comparatively lesser utility.

There are two take-aways from these highly variable vowel data. First, we see that all of these features display variability in baseline and range across vowels, suggesting that across-vowel comparison of raw measurement values, even within a speaker, may not yield meaningful results. Second, we see that the high variability and low ΔFeature values indicate that A1-P1 was not a useful feature in these data, and suggest more reliable results may be obtained simply using A1-P0 to measure nasality in all vowels.

Finally, we must again state that the features here are meant to be vowel-general, the raw acoustic consequence of adding nasal coupling to any vowel. Given that nasal vowels are also articulated with specific patterns, measuring and examining these language- and vowel-specific oral articulations may provide a separate and useful set of features for nasality, which may be helpful both in measurement and in perception of this complex phenomenon.

C. Speaker variability

Figure 3 shows the oral and nasal mean for each speaker in the in English and French datasets, again using the most promising acoustic features. Here, we see the grand average across all vowels for each feature in both oral and nasal vowels in the dataset. For instance, we see that speaker en10 had a mean A1-P0 of ∼5 in oral vowels, and ∼0 in nasal vowels.

First, we see that the speakers vary greatly in terms of A1-P0, both in their oral and nasal baseline measurements, as well as in their range of A1-P0 from oral and nasal vowels, indicating that a change of n dB A1-P0 is not a uniform shift in degree of nasalization across speakers. Even where the range is similar across two speakers, we see pairs, such as with speakers en5 and en6 in the English data, where one speaker's oral-vowel mean A1-P0 is nearly identical to the other's "nasal" mean. Thus, a speaker showing lower raw values for A1-P0 cannot reliably not be judged as "more nasal" than another based on A1-P0 alone.

For A1-P1, in addition to similar variability in range and baseline, we see that some speakers simply do not show an A1-P1 effect (en8 in English, as well as fr3, fr4, fr6, and fr7 in French). This is consistent with prior observations that A1-P1 is variably measurable across speakers, and the comparison with A1-P0 again shows that A1-P0 is more reliable, particularly for semi-automated or automated measurement.

We see that spectral tilt (A3-P0) again changes in the predicted direction for all speakers, as well as baseline differences in spectral tilt. Interestingly, the across-speaker variation in range is relatively small among speakers of each language, but the difference between the French and English speakers is rather large, with English speakers uniformly showing a far smaller ΔA3-P0 than French speakers. This reflects the findings from the model, where French speakers showed greater separation in these features between oral and nasal vowels.

Finally, we see that F1's bandwidth shows great variability in baseline and range, which is to be expected from a formant-dependent feature, but again, as with spectral tilt, the English speakers largely show less ΔBandwidth than the French speakers, again reflecting the findings of the model.

We find, then, that speaker variability for all of the measures is widespread, even across an identical set of words, and none of our measurements are interpretable across-speakers without some form of normalization. This is, at some level, unsurprising, given that speakers will differ both in the nasal and oral anatomy which give rise to the resonances associated with nasality, as well as exhibiting the well-attested across-speaker differences in vowel space and formant patterns, which nasal resonances then interact with. When coupled with the fact that three of these four measures are based directly on measures of the vowel formants, the across-speaker variability in these nasality measurements seems likely to be as complex, if not more so, than we already face in comparing vowel quality measurements.

Thus, much like absolute values for vowel formants, direct comparison of acoustic nasality measurement values across speakers does not provide meaningful information about degree (or even presence) of nasality, and at the moment, there is no reliable means of directly comparing speakers' degree of nasality based on acoustic data alone. Although continued work on aerodynamic and physiological approaches may help to address this problem in some research contexts, this suggests a need for the development of algorithmic normalization methods for acoustic nasality measurements, akin to Lobanov or Nearey transformations of vowel formants, to allow these across-speaker comparisons to be more reasonably made. It also suggests that even for within-speaker comparisons, researchers should attend to both the differences in baseline and range, centering A1-P0 values as well as directly measuring the oral-to-maximally nasal range using oral (in English, CVC) and maximally nasal (in English, NVN) words.

Thus, although across-vowel variability must be kept in mind, and we appear to face a similar speaker normalization problem in comparing raw nasality measures as we do with comparing vowel formants, when used carefully, A1-P0, as described in Chen 1997, remains an especially useful feature for the acoustical measurement of nasality.

D. The acoustics of nasality in English vs French

Given these findings, we can now address our final question: Do English and French differ not just in the phonological nature of nasality, but in its acoustic expression as well?

At some level, the acoustics consequences of nasality in these two languages are fairly similar. All but four of the features that showed a correlation with nasality in English showed a similar correlation in French and some of that variation, particularly the lack of an A1-P1 or A1-P2 effect in French, was likely due to features showing effects only in certain vowels, absent in the French dataset. In addition, for every acoustic feature that showed a significant correlation with nasality in both languages, the direction of ΔFeature was the same, and no acoustic phenomenon showed a correlation in either language which was completely unattested in any form in the other. However, even with these similarities, there is some evidence that nasality in English is, acoustically speaking, meaningfully different from nasality in French.

Figure 4 shows across-speaker, by-vowel means for the closest equivalent vowels in both English and French, for the three features which were significant in both languages. For A1-P0, although the difference in raw ΔA1-P0 is hard to interpret directly due to variability in range and baseline, we see roughly comparable patterns, albeit with higher values in French /ɔ/. However, looking at the effect size, shows greater evidence of difference: For A1-P0, French shows a CoefVsSD of 1.105, relative to the English value of 0.563, indicating that French speakers produce nasality in such a way that it is *far* more distinctive relative to the noise in oral vowels. If this difference is attributable to language-specific differences in degree of nasality, we would expect similar gains in CoefvsSD in all nasality-associated measures. However, if we look at effect size for F1's bandwidth, we see a very different relationship, with French CoefVsSD for formant bandwidth at 0.43, versus the roughly similar 0.478 in English. This seems to suggest that French speakers are producing not just a greater degree of nasality, but a stronger A1-P0, relative to the variation in oral vowels.

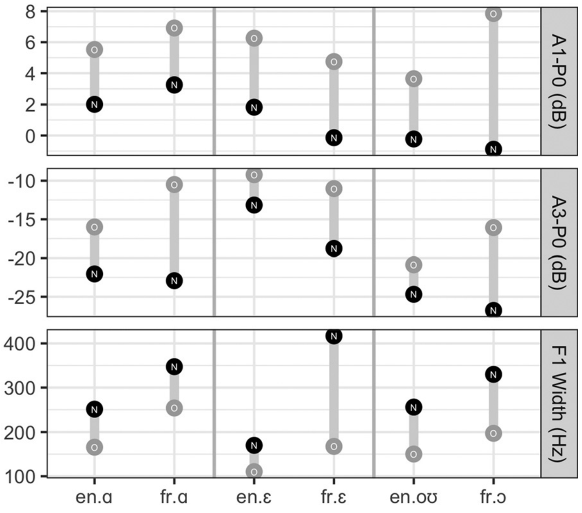

FIG. 4. Oral (gray) and nasal (black) means by feature, in English (marked "en") and French (marked "fr") vowels. Features scaled individually.

In F1's bandwidth, although we see greater ΔBandwidth in /ɛ/ in French, the differences in the other vowels are not substantial. This rough equivalence is supported by the model output (Tables V and VI), which show slight differences in ΔBandwidth (∼157 Hz) for French relative to English, where we saw ΔBandwidth of ∼96 Hz. That said, the two languages show no substantial difference in CoefvsSD (0.478 for English, 0.432 for French), implying that bandwidth is not systematically more distinctive for nasality in either

For spectral tilt, the difference between English and French is quite apparent. In English, all vowel qualities and speakers uniformly showed smaller changes in spectral tilt in nasal vowels relative to oral vowels. In French, though, ΔTilt was *far* stronger, both in terms of CoefVsSD and in terms of absolute numbers. The strength and uniformity of this effect among the French speakers strongly implies that, indeed, French speakers are producing nasal vowels in a meaningfully different way, possibly enhancing the natural tendency of nasal pole-zero complexes towards increased spectral tilt, perhaps through a breathier voice quality. This is in line with the enhancement of nasality through breathy voicing which was documented in Bo, Luchun Hani, and Southern Yi in Ref. 26.

Based on these findings, it does appear that the French speakers are creating vowel nasality which is not just differently timed or greater in degree, but acoustically different *and more functionally distinct* than English nasality.

That French speakers might heighten the oral-nasal contrast is perhaps unsurprising. Given its phonemic status, vowel nasality in French carries a considerably higher functional load than in English, where a disambiguating consonant is often present (although not always, cf. Ref. 41). Thus, we might expect speakers to (subconsciously) produce nasality in a way which would increase the oral-to-nasal separation for meaningful features, and thus, optimize the discriminability of nasality for listeners, whether through shifts in oral articulation (as described in Ref. 17, among others), by increasing the degree of nasalization, or by enhancing already-present acoustical features in the signal beyond the level naturally expected. It is also possible, in the case of breathy voicing and spectral tilt, that French speakers are recruiting other features and articulations to further enhance the contrasts. Thus, we must be open to nasal vowels having complex realizations, far beyond "oral vowels with a lowered velum."

So, although there were more similarities than differences in the overall implementation of nasality in English and French, given the sharp difference in spectral tilt, variation in the signal-to-oral-noise, and differences in degree for many of our features, these data imply that the acoustics of nasality do differ in these two languages. Thus, nasality should be considered language-specific, not just in timing and degree, but in the nature of the changes and features used to maintain their distinctiveness.

VII. CONCLUSIONS

In the present study, we collected two corpora of elicited speech data, one in English and one in French, then measured and analyzed 22 different acoustical features in both languages. From these data, we can draw three main conclusions.

First, we have a clearer understanding of which acoustical features seem to result from the addition of nasal coupling and their relative reliability, pointing to A1-P0, F1's bandwidth and spectral tilt as the most robust vowel-general features of nasality in French and English, with A1-P0 appearing most reliable for use as a measurement of nasality.

Second, we see that these acoustic features, particularly A1-P0, show significant variability both in baseline and oral-to-nasal range across speakers, highlighting that, much like vowel formants, acoustical nasality measurements are highly variable across speakers. This highlights a need for future work to develop a normalization algorithm for these measures which will reduce the impact of this variability. But more importantly, it shows the need for these measures, much like vowel formants, to be treated as "unnormalized," and thus, not directly comparable across speakers.

Finally, we find that although vowel nasality produces similar acoustic consequences in both French and English, there are notable differences, both in the degree of nasality and in the oral-nasal contrasts features used. Particularly, this study shows a sharp increase in spectral tilt in French nasal vowels, suggesting that these French speakers are using breathy voice to enhance the already-present nasal spectral tilt in order to heighten the contrastiveness of nasality in speech. This suggests that although the same basic patterns may be found across languages, the exact acoustical nature of nasality is language-specific, and urges us towards further research linking nasal acoustics and nasality's functional load in other languages.

Future work from this project will discuss subsequent experiments probing both the predictive power of these features for machines, and the perceptual utility of these features for humans. But for now, the existence of these three useful features, albeit with considerable across-language, across-vowel, and across-speaker variability, highlight the deep, but tractable, acoustical complexity of vowel nasalization.

ACKNOWLEDGMENTS

The author would like to thank Rebecca Scarborough, Pam Beddor, Andries Coetzee, and Jelena Krivokapic for their guidance, input, support, and overall kindness. The author would also like to thank Luciana Marques for recording the French speakers in Colorado, and Georgia Zellou for recording the French speakers at Penn, as well as Kerby Shedden for his input on the statistical methods used, and two anonymous reviewers for their valuable feedback. The author would also like to thank Martha Palmer, Kathy Arehart, Mans Hulden, David Rood, Wayne Ward, John Ohala, Jacqueline Vaissiere, Luciana Marques, and Lise Menn for their additional input, insights, and pleasant conversations. Finally, my sincere gratitude to the twenty participants whose hard velopharyngeal work, through real words and non, made this study possible. This work is supported in part by NSF Grant No. BCS-1348150 to Patrice Speeter Beddor and Andries W. Coetzee.

[1] M. Haspelmath, *The World Atlas of Language Structures* (Oxford University Press, Oxford, 2005).
[2] A. Lahiri and W. Marslen-Wilson, "The mental representation of lexical form: A phonological approach to the recognition lexicon," Cognition **38**, 245–294 (1991).
[3] P. S. Beddor and R. Krakow, "Perception of coarticulatory nasalization by speakers of English and Thai: Evidence for partial compensation," J. Acoust. Soc. Am. **106**, 2868–2887 (1999).
[4] R. Scarborough, "Neighborhood-conditioned patterns in phonetic detail: Relating coarticulation and hyperarticulation," J. Phon. **41**, 491–508 (2013).
[5] R. Scarborough, "Lexical similarity and speech production: Neighborhoods for nonwords," Lingua **122**, 164–176 (2012).
[6] W. Styler, "On the acoustical and perceptual features of vowel nasality," Ph.D. dissertation, University of Colorado at Boulder, Boulder, CO (2015).
[7] M. Y. Chen, "Acoustic correlates of English and French nasalized vowels," J. Acoust. Soc. Am. **102**, 2360–2370 (1997).
[8] M. Y. Chen, "Acoustic parameters of nasalized vowels in hearing-impaired and normal-hearing speakers," J. Acoust. Soc. Am. **98**, 2443–2453 (1995).
[9] M. Schwartz, "The acoustics of normal and nasal vowel production," Cleft Palate J. **5**, 125–140 (1968).
[10] K. N. Stevens, *Acoustic Phonetics* (MIT Press, Cambridge, MA, 1998), pp. 303–322.
[11] V. Delvaux, T. Metens, and A. Soquet, "Propriétés acoustiques et articulatoires des voyelles nasales du Français" ("Acoustic and articulatory properties of French nasal vowels"), in *XXIVèmes Journées d'Étude sur la Parole*, Nancy (June 24–27, 2002), Vol. 24, pp. 357–360.
[12] N. Macmillin, J. Kingston, R. Thorburn, L. W. Dickey, and C. Bartels, "Integrality of nasalization and F1. II. Basic sensitivity and phonetic labeling measure distinct sensory and decision-rule interactions," J. Acoust. Soc. Am. **106**, 2913–2932 (1999).
[13] T. Pruthi and C. Espy-Wilson, "Acoustic parameters for automatic detection of nasal manner," Speech Commun. **43**, 225–239 (2004).

[14] V. Delvaux, “Perception du contraste de nasalité vocalique en Français” (“Perception of vocalic nasal contrast in French”), J. French Lang. Stud. **19**, 25–59 (2009).

[15] R. A. Krakow, P. S. Beddor, L. M. Goldstein, and C. A. Fowler, “Coarticulatory influences on the perceived height of nasal vowels,” J. Acoust. Soc. Am. **83**, 1146–1158 (1988).

[16] C. Carignan, “An acoustic and articulatory examination of the oral in nasal: The oral articulations of French nasal vowels are not arbitrary,” J. Phon. **46**, 23–33 (2014).

[17] C. Carignan, R. Shosted, M. Fu, Z. Liang, and B. P. Sutton, “A real-time MRI investigation of the role of lingual and pharyngeal articulation in the production of the nasal vowel system of French,” J. Phon. **50**, 34–51 (2015).

[18] R. Shosted, C. Carignan, and P. Rong, “Managing the distinctiveness of phonemic nasal vowels: Articulatory evidence from Hindi,” J. Acoust. Soc. Am. **131**, 455–465 (2012).

[19] C. Carignan, R. Shosted, C. Shih, and P. Rong, “Compensatory articulation in American English nasalized vowels,” J. Phon. **39**, 668–682 (2011).

[20] J. Kingston, in *The Cambridge Handbook of Phonology*, edited by P. De Lacy (Cambridge University Press, Cambridge, UK, 2007), pp. 435–456.

[21] V. Delvaux and K. Huet, “Perception de la nasalité en Français de Belgique: Catégorisation dirigée et catégorisation libre” (“Perception of nasality in French from Belgium: Controlled categorization and free categorization”), Rev. Parole **3**, 137–176 (2006).

[22] S. Hawkins and K. N. Stevens, “Acoustic and perceptual correlates of the non-nasal–nasal distinction for vowels,” J. Acoust. Soc. Am. **77**, 1560–1575 (1985).

[23] S. Maeda, “Acoustics of vowel nasalization and articulatory shifts in French nasal vowels,” Phon. Phonol. **5**, 147–167 (1993).

[24] T. Arai, “Cue parsing between nasality and breathiness in speech perception,” Acoust. Sci. Technol. **27**, 298–301 (2006).

[25] A. P. Simpson, “The first and second harmonics should not be used to measure breathiness in male and female voices,” J. Phon. **40**, 477–490 (2012).

[26] M. Garellek, A. Ritchart, and J. Kuang, “Breathy voice during nasality: A cross-linguistic study,” J. Phon. **59**, 110–121 (2016).

[27] K. Stevens and H. Hanson, “Classification of glottal vibration from acoustic measurements,” in *Vocal Fold Physiology: Voice Quality Control*, edited by O. Fujimura and M. Hirano (Singular Pub. Group, Buckinghamshire, UK, 1995), pp. 147–170.

[28] M. Iseli, Y. Shue, and A. Alwan, “Age, sex, and vowel dependencies of acoustic measures related to the voice source,” J. Acoust. Soc. Am. **121**, 2283–2295 (2007).

[29] J. Yuan and M. Liberman, “Speaker identification on the SCOTUS corpus,” in *Proceedings of Acoustics 2008*, Paris, France (2008), pp. 3878–3878.

[30] P. Milne, “The variable pronunciations of word-final consonant clusters in a force aligned corpus of spoken French,” Ph.D. dissertation, University of Ottowa, Ottowa, ON (2014).

[31] J. Ohala and M. Amador, “Spontaneous nasalization,” J. Acoust. Soc. Am. **69**, S54 (1981).

[32] P. Boersma and D. Weenink, “Praat: Doing phonetics by computer” [computer program] (Version 6.0.28), http://www.praat.org (Last viewed June 29, 2017).

[33] https://github.com/stylerw/styler_praat_scripts/tree/master/nasality_automeasure (Last viewed October 19, 2017).

[34] A. Cohn, “Phonetic and phonological rules of nasalization,” UCLA Working Papers in Linguistics, UCLA Department of Linguistics, Los Angeles, CA (1990), Vol. 76.

[35] D. Bates, M. Maechler, B. Bolker, and S. Walker, *lme4: Linear mixed-effects models using Eigen and S4* (2014), R package version 1.1-6.

[36] R. Baayen, *Analyzing Linguistic Data: A Practical Introduction to Statistics Using R* (Cambridge University Press, Cambridge, 2008), pp. 247–248.

[37] J. Laver, *The Phonetic Description of Voice Quality* (Cambridge University Press, Cambridge, England, 1980), pp. 132–135.

[38] M. Gordon and P. Ladefoged, “Phonation types: A cross-linguistic overview,” J. Phon. **29**(4), 383–406 (2001).

[39] A. Sluijter and V. Van Heuven, “Spectral balance as an acoustic correlate of linguistic stress,” J. Acoust. Soc. Am. **100**, 2471–2485 (1996).

[40] P. Murphy, K. McGuigan, M. Walsh, and M. Colreavy, “Investigation of a glottal related harmonics-to-noise ratio and spectral tilt as indicators of glottal noise in synthesized and human voice signals,” J. Acoust. Soc. Am. **123**, 1642–1652 (2008).

[41] P. S. Beddor, “A coarticulatory path to sound change,” Language **85**, 785–821 (2009).