

Chapter 10

Revisiting sociophonetic competence: Variable spectral moments in phrase-final fricative epithesis for L1 & L2 speakers of French

Amanda Dalola^a & Keiko Bridwell^b


^aUniversity of Minnesota ^bUniversity of Georgia

Phrase-final fricative epithesis (PFFE) is a phenomenon in Continental French in which utterance-final vowels lose their voicing and yield fricative-like whistles corresponding to the identity of the host vowel. PFFE is also well attested among L2 speakers; differences in its production across speaker groups have been reported in several domains: vowel type, speech rate, register, constituent location, fricative-vowel ratio (FVR), and measures of center of gravity (COG). Participants completed a reading task targeting 98 tokens of /i,y,u/ in phrase-final position. 4569 PFFE segments were examined via 7 frames of 8 ms in length with 2 ms in overlap across their trajectories. Results suggest categorical speaker group differences for /y/ in terms of skewness, as well as for kurtosis and for intensity at higher FVRs. This suggests that L1/L2 sociophonetic realizations contain nuanced differences far beyond the presence/absence paradigm still common to many variationist inquiries.

1 Introduction

Phrase-final fricative epithesis (PFFE), a phenomenon also known in the literature as phrase-final vowel devoicing (PFVD), refers to a well-attested phenomenon in Continental French (CF) in which breath group-final vowels lose their voicing and produce a short burst of high-frequency aperiodic energy akin to a



Amanda Dalola & Keiko Bridwell. 2023. Revisiting sociophonetic competence: Variable spectral moments in phrase-final fricative epithesis for L1 & L2 speakers of French. In Barbara E. Bullock, Cinzia Russi & Almeida Jacqueline Toribio (eds.), *A half century of Romance linguistics: Selected proceedings of the 50th Linguistic Symposium on Romance Languages*, 195–220. Berlin: Language Science Press. DOI: 10.5281/zenodo.7525110 

fricative, e.g. *mais oui-hhh* [mɛwiç], *merci beaucoup-hhh* [mɛʁsibokux] (see Figure 1). The first linguistic description of this phenomenon described it as the emergence of “sharp, phrase-final whistles” (Fónagy 1989); subsequent research witnessed a split in nomenclature, with North American researchers often opting for a name focusing on voicing loss—“vowel devoicing / *dévoisement vocalique*” (Fagyal & Moisset 1999, Smith 2002, 2003, 2006, Martin 2004) and most European researchers preferring a name focusing on the emergence of the downstream fricative – “fricative epithesis / *épithèse (consonantique) fricative*” (Fagyal 2010, Candea 2012, Candea et al. 2013). Because the present study will focus on characterizing the spectral and durational qualities of the emergent fricative, we, the (North American) authors, have explicitly chosen to heed the call of our European predecessors in adopting the term “fricative epithesis” for this discussion, a precedent we first established in Dalola & Bridwell (2020).

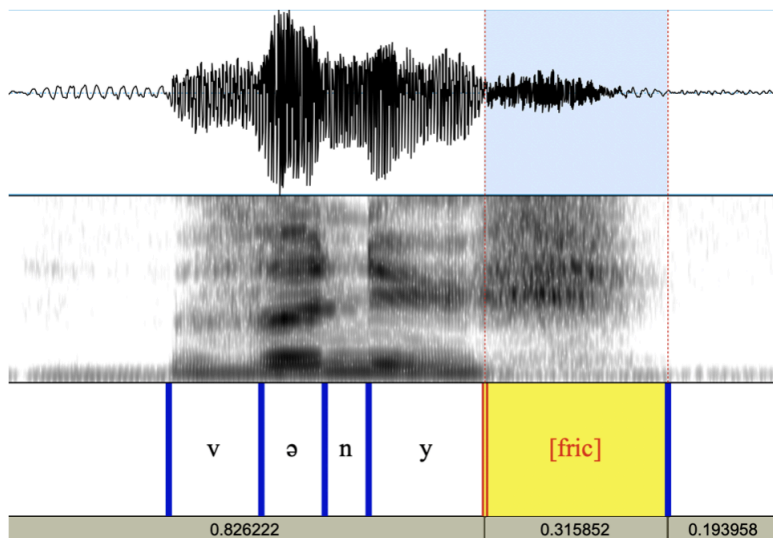


Figure 1: PFFE on the spectrogram: *venu* ‘came’. The PFFE corresponds to the final, highlighted segment, which is characterized by the lack of a voicing band on the spectrogram and the presence of aperiodic energy in the waveform. The PFFE is immediately following the articulation of the vowel [y], which is characterized by full formant structure on the spectrogram and periodic energy in the waveform.

In the first description of PFFE in the literature, Fónagy (1989) hypothesized that not only do the characteristic phrase-final fricatives appear immediately following vowels that have lost a portion of their voicing band, but they themselves might also correspond to their host vowel phonetically in terms of their backness

dimension. Citing the *ich-Laut* / *ach-Laut* harmony phenomenon in standard German, in which the backness value of a voiceless fricative is selected by the backness value of its preceding vowel, he hypothesized that the fricatives epithesized after the high front vowels /i/ and /y/ in French should be more [ç]-like, i.e. front, than those appearing after high back [u], which should be more [x]-like, i.e. back. This observation was corroborated by Dalola (2015a) who examined measures of center of gravity (COG: average peak frequency) taken at the 1/4, 1/2 and 3/4 timepoints of PFFE fricatives produced by L1 CF speakers and found evidence to suggest a three-way distinction in spectral energy at the first two timepoints. These spectral differences, however, could not be characterized in terms of sheer [+/- back] and did not persist into the second half of the segment.

1.1 Phonological Predictors of PFFE

The best-studied dimension of PFFE is undoubtedly its phonological distribution. Originally described as occurring in high vowels (Fónagy 1989), PFFE has been documented in the full inventory of French vowels, including nasals (Smith 2006), but has been reported at the highest rates following the high vowels /i,y,u/ (Fagyal & Moisset 1999, Martin 2004, Smith 2003, 2006). When comparing reading passages, role-plays and impromptu conversation, PFFE has been found to occur at significantly higher rates in types of reading, i.e. planned, speech (Fagyal & Moisset 1999, Dalola 2014). This finding is perhaps explained by its higher rates of occurrence at the ends of both the intonation phrase and the declarative phrase (Fagyal & Moisset 1999, Smith 2003) where French sees the emergence of a low tone. Studies have also found an effect for the manner type of the preceding consonant, such that preceding stops condition PFFE at a significantly higher rate than more sonorous manner types, in addition to lexical frequency effects, which report more frequent lexical items as more likely to exhibit the phenomenon than less frequent ones (Dalola 2015b). Dalola & Bridwell (2019, 2020) also found that PFFE varies with the proportional duration of the fricative relative to the full vowel (i.e. FVR, as described in §2.4), such that among high front vowels, longer fricatives are produced with higher COGs, indicating that epithesized fricatives may be conditioned by the same principles of hypo- and hyperarticulation (Lindblom 1990) as other phonological segments.

1.2 Social Predictors of PFFE

The social distribution of PFFE presents a complex series of macro- and micro-group associations. Early work often described PFFE as occurring in the speech

of women (Fónagy 1989, Fagyal & Moisset 1999, Smith 2006); however, later work has reported the variable to be used at similar rates among both men and women (Candea 2012, Candea et al. 2013, Dalola 2014). Fagyal & Moisset (1999), who took a categorical approach to age, found the variable at its highest rates among their youngest (16–35) and oldest (61–85) groups; Dalola (2014), who operationalized age continuously (testing ages 13–83), reported participants as more likely to use PFFE the older they were. From a socioeconomic standpoint, PFFE is often associated with the French middle class (*la bourgeoisie*) (Paternostro 2008, Fagyal 2010). Originally, the variable was associated with Parisians (Fagyal & Moisset 1999, Smith 2006, Fagyal 2010), though in recent years, it has also been documented in the speech of francophones from other metropolitan centers in France, namely Lyon and Strasbourg (Dalola 2014). Further afield, the variable has also been described in the speech of French, Belgian, and Canadian news anchors (Paternostro 2008, Candea et al. 2013); one study introduced intersectionality into this association by reporting it as a characteristic of young, i.e. inexperienced, news anchors (Candea 2012). The disagreement among social predictors may be reflective of the rapidly changing landscape of PFFE users and/or the generalization of the variable from specialized demographics to less specialized ones. Despite the inconclusive findings among social predictors, it is important to recognize them in pursuing research on the characterization of the PFFE variable.

1.3 Perception of PFFE

Differences in L1 versus L2 production of PFFE ushered in a rigorous examination of potential speaker group differences in the variable's perception. Dalola (2016) reported significant differences in L1 and L2 perceptions of the PFFE, namely that L2 speakers perceived it as a marker of "formality" and "trustworthiness," whereas L1 speakers perceived it variably, sometimes as a marker of "admirability" and other times as a marker of "intense emotional affect." Using a matched-guise design and exploratory factor analysis, a related form of principal component analysis that partitions out the shared variance of each variable from its unique and error variance to reveal the underlying factor structure (Osborne & Costello 2009), Dalola's (2016) L2 participants rated users of PFFE similarly for two separate groups of adjectives: *polite, well-educated, speaks clearly, speaks formally*, a category the author refers to collectively as traits of FORMALITY, and *confident, persuasive, I respect X, I trust X*, a category the author refers to collectively as traits of TRUSTWORTHINESS. Like their L2 counterparts, L1 speakers also rated users of PFFE similarly for two separate groups of adjectives, however, the

adjectival members of the groups were both more numerous and compositionally different: *well-educated, professional, speaks clearly, polite, intelligent, patient, confident, persuasive, I trust X, I respect X, I believe what X says, I would like to speak like X* formed a category referred to collectively as traits of ADMIRABILITY, while *aggressive, bourgeois, superficial, bossy, native French speaker, speaks with emotion* formed a category referred to collectively as EMOTIONAL AFFECT. It should be noted that all the traits that made up the L2 category of FORMALITY were also present in the L1 category of ADMIRABILITY, and that the reason for the difference in category name was due to the author's desire to assign names that applied to the full collection of adjectives. No gender effects were found for the voices being rated, however, there was a significant gender effect among those giving ratings, such that women were more likely to assign higher ratings for the adjective profiles overall than men.

1.4 L2 Speakers & PFFE

Given its salient phonetic energy and robust distribution among native franco-phone populations, it is somewhat unsurprising to learn that PFFE, despite its status as a sociophonetic variable, is also readily employed by L2 French speakers (Dalola & Bullock 2017). Investigating the nature of L1 and L2 PFFE as produced in different genres of speech, they revealed subtle but nuanced differences at every level of production. For rates of use of PFFE, L1 and L2 speakers performed similarly overall but were motivated by different genres of speech: L1 speakers used more PFFE in role-plays while L2s were more likely to use it when reading wordlists. In terms of PFFE duration, or the proportional length of the epthesized fricative when compared to its host vowel, larger differences between speaker groups were documented: not only did L1 and L2 speakers produce PFFE segments that were statistically different in length (L1 PFFE length > L2 PFFE length), but each group showed sensitivity to a different linguistic parameter. L1s produced longer PFFEs as a reaction to pragmatic shifts (indicated in the role-plays via scenario-setting prompts), producing longer PFFEs in slower and formal speech, while L2s produced longer PFFEs as a reaction to task shifts (indicated by shifting from roleplays to wordlist), producing longer PFFEs in the wordlist task. Despite the various pragmatic and speaker group effects in this study, no effects were found across participants for measures of gender or age. In the first study to examine spectral differences between L1 and L2 users of PFFE, Dalola & Bridwell (2020) examined COG measures taken throughout the fricative and found L1 speakers to produce PFFE differently, as a function of phonological context,

while L2 speakers instead were found to produce a singular, high-energy, hyper-articulated allophone in a majority of environments. The discovery of spectral differences in terms of COG measures, a metric commonly used for diagnosing fricatives (segments that are notoriously difficult to describe due in part to their sustained aperiodic energy), therefore, invites the investigation of the other spectral moments commonly used to talk about fricatives, namely standard deviation, skewness, and kurtosis, as well as independent measures of intensity.

1.5 Spectral Moments

The use of the four spectral moments in diagnosing the energy of fricatives is common practice in phonetic studies, as it has been reported that listeners tend to classify voiceless obstruents from spectral information during the first 40 ms of the segment (Forrest et al. 1988). The four spectral moments used to classify fricatives are individually center of gravity (COG), standard deviation (SD), skewness, and kurtosis; each is a unique statistical manipulation of the segment's energy profile meant to capture a different quantitative aspect of the aperiodic energy sustained throughout the segment. If one conceives of the various frequencies produced during a fricative in terms of a normal distribution (see Figure 2), then the COG corresponds to the measure of central tendency among the frequencies, or the spectral mean. The standard deviation is then the measure of spread, or the average variation in frequencies from the spectral mean during the segment. As with distributions, skewness corresponds to a measure of energy distribution above or below the spectral mean, or spectral tilt, while kurtosis corresponds to a measure of the distribution of energy at the tail extremities, or tailedness. In addition to the four spectral moments, this study will also take measurements of amplitude, also called intensity, which is perceived by listeners as loudness. Although not one of the four spectral moments, intensity has also been shown to be a relevant parameter for the perception of obstruents (Forrest et al. 1988).

1.6 Spectral moments as cues

Although spectral moments and measures of intensity have been shown to aid listeners in identifying and discerning voiceless obstruents in languages, they have also been found working independently or in concert with other acoustic parameters to convey social information about the speaker. Zimman (2017) reported COG to be a marker of masculine voices when considered alongside fundamental frequency (f_0). Munson et al. (2006) found the skewness of frequencies in [s] to be a relevant predictor of speaker sexuality, such that negatively skewed [s],

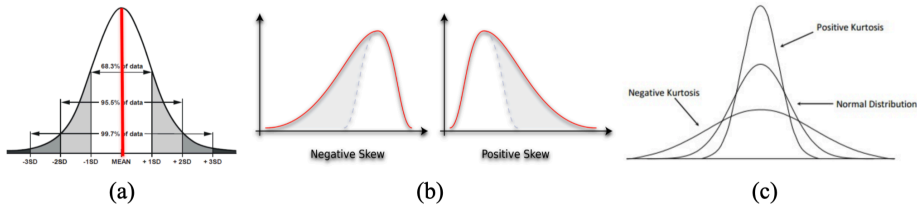


Figure 2: The four spectral moments, visualized as statistical transformations: a) center of gravity and standard deviation, b) skewness, c) kurtosis.

or [s] segments with a higher concentration of energy above the spectral mean, served as a marker of gay men's speech. In terms of the PFFE variable, Dalola (2016) found intensity measures to be a predictor of affect ratings assigned by L1 French listeners to PFFE users, such that higher intensity PFFEs yielded higher ratings for features of INTENSE AFFECT (*aggressive, bourgeois, superficial, bossy, native French speaker, speaks with emotion*). Although relevant for the perception of INTENSE AFFECT, intensity measures were not found to be a predictor of features of ADMIRABILITY (*educated, professional, speaks clearly, polite, intelligent, patient, confident, persuasive, I trust X, I respect X, I believe what X says, I would like to speak like X*) among L1 French listeners, nor were they found to be at all relevant for L2 French listeners perceiving PFFE segments.

1.7 Motivation

This article reports on production differences in spectral tendencies in PFFE among L1 and advanced L2 speakers of Continental French. Since PFFE is a sociophonetic marker in CF (Dalola 2014, 2016), it presents an interesting testing ground for comparing spectral values across native and non-native speakers. Previous work has reported production differences in rate and degree of devoicing between native and non-native French speakers (Dalola & Bullock 2017), but only a few studies have extended the comparison to investigate the phonetic quality of PFFE's variable emergent fricatives (Dalola & Bridwell 2020). This study looked uniquely at COG measures and reported significant differences across speaker groups. However, fricatives are most commonly identified by a profile of four spectral moments, including measures of COG, SD, skewness, and kurtosis, not just COG alone. Combined with the many known articulatory differences and false similarities in vowel production between French and English (the L1 of the non-native population in this/previous studies), e.g. French /u/=[u] vs. English

/u/=[u^w], it is reasonable to expect that L1 articulatory behaviors may persist which could contribute to differences in spectral energy when realizing PFFE, even among advanced L2 speakers (Flege & Hillenbrand 1984, Flege 1987, 1995). The goal of this study, therefore, is to examine and characterize the fricatives epithesized after devoiced vowels using measures of fricative-vowel ratio (length of fricative divided by length of full vowel), COG (average peak frequency during PFFE segments), SD (variation in frequencies from the spectral mean during PFFE segments), skewness (energy distribution above or below the spectral mean during PFFE segments), kurtosis (energy distribution at the tail extremities during PFFE segments), and intensity (loudness of PFFE segments). We will then use inferential statistics to determine the presence of speaker group differences in PFFE in terms of measures of COG, SD, skewness, kurtosis, and intensity. Finally, we will investigate the role of vowel type and fricative-vowel ratio in modulating these speaker group differences.

1.8 Research Questions & Hypotheses

The current study puts forth the following research questions:

1. Do L1 and advanced L2 French speakers produce PFFE differently in terms of measures of COG, SD, skewness, kurtosis or intensity?
2. Are speaker group-level differences in the production of PFFE modulated by vowel type and fricative-vowel ratio?

A previous study carried out by Dalola & Bridwell (2020) predicts significant differences across speaker group for measures of COG, a hypothesis that is maintained in the present study. Predictions will not be offered for each of the other parameters, as previous work has not yet diagnosed these aspects of the PFFE variable among L1 or L2 speakers.

2 Methods

2.1 Participants

40 speakers of CF participated in the experiment, of which 31 were L1-French speakers and nine L1-English advanced L2-French speakers. All participants were recorded in Paris or Strasbourg in France or in the United States. Among the L1 participants, 23 were women and eight were men, ranging in age from 20 to 66 years (mean: 38.4 years). All L1 speakers were L2 speakers of English, having

studied it formally for four or more years and using it in interactions once a week or more. Among the L2 participants, five were women and four were men, ranging in age from 27 to 58 years (mean: 38.6 years). L2 speakers were classified as “advanced” because they had all lived in France for at least two years, had prepared or were preparing an upper-level degree in French, and used French regularly in their careers. All L2 speakers were L1 speakers of American English. All subjects gave their informed consent for inclusion before they participated in the study. The study was conducted in accordance with the Declaration of Helsinki, and the protocol was approved by the University of South Carolina Institutional Review Board (USC IRB).

2.2 Stimuli

Inspired by the task and pragmatic effects findings of Dalola & Bullock (2017) and several studies’ reports of PFFE’s robustness among news anchors reading off teleprompters (Paternostro 2008, Candea 2012, Candea et al. 2013), participants were asked to complete a reading task that consisted of 106 single sentences, containing 98 phrase-final tokens of /i,y,u/, occurring after all licit (C)C(C) onset sequences in one- to three-syllable real words in French (see Table 1 for a breakdown of consonant environments).

Table 1: Stimuli consonant onset sequences

Number of tokens	Consonant onset sequences
3	/p, t, k, b, d, g, f, s, ʃ(S), v, z, ʒ(Z), r, l, m, n, kr, tr, bl, gr, fl/
2	/pl, pr, gl, dr, fr, ʒ(Z)r, ʃ(S)r/
1	/ɥ(4), j, w, kl, bl, pɥ(4), tɥ(4), kɥ(4), bɥ(4), dɥ(4), fɥ(4), sɥ(4), nɥ(4), lɥ(4), lw, bw, fw, lw, nw, plɥ(4), trɥ(4)/

2.3 Procedure

Participants were presented with sentences one at a time on a MacBook Pro via Microsoft Powerpoint and told to read each one aloud, imagining they were reading a story to a native francophone listener. Participants were instructed to read each sentence twice and to repeat any trials from the beginning in the event of a disfluency. As they read aloud, participants were recorded via a head-mounted unidirectional cardioid microphone (SHURE WH20) plugged into a solid-state

digital recorder (Marantz PMD 660) digitized at 44.1 kHz (16 bit). The task was completed under the direction of the L1-English advanced L2-French researcher; the task was self-paced and participants were given as much time as they needed to complete it.

2.4 Measurements

The final vowel of each target word was examined in Praat (Boersma & Weenink 2018) for the presence of PFFE, as indicated by the loss of a voicing band and the onset of high-frequency aperiodic energy (for reference, see Figure 1). All tokens exhibiting PFFE were then measured for the duration of the fricative and of the vowel itself. Since measures of raw duration do not take into account relative durational differences brought on by variable realizations of final lengthening, the derived measure of *fricative-vowel ratio* (*FVR*) was selected as the relevant durational variable. This measure was calculated according to the formula $FVR = F/V$, where F is the length of the epithesized fricative and V is the length of the full vowel including the fricative (Dalola & Bridwell 2020).

Time-averaged spectral measures were obtained for each of the tokens by measuring the spectra of seven frames, each 8 ms in length and overlapping by 2 ms, from the central 44 ms of the fricative and averaging across these frames (see Toda 2007, 2009). This method of averaging across multiple windows was chosen in order to reduce the effects of the random fluctuation that is inherent in aperiodic energy (Toda et al. 2010). The resulting measures included the four spectral moments of *COG*, *standard deviation* (*SD*), *skewness*, and *kurtosis*, and a measure of *intensity*.

2.5 Statistical Analyses

Of the 7942 tokens collected, 4995 exhibited PFFE (62.9%). Of these, 426 (8.5%) were discarded as being too short (i.e. exhibiting under 44 ms of frication) or as spectrally unanalyzable at one or more of the timepoints, leaving 4569 tokens for analysis. Statistical analyses were conducted in R (R Core Team 2018), using the function *lmer()* from the packages *lme4* (Bates et al. 2015) and *lmerTest* (Kuznetsova et al. 2017) to run mixed-effects linear regression models for each of the five variables. This method of analysis was chosen because mixed-effects models allow for reliable comparisons between unequal sample sizes. In each model, *vowel*, *speaker group*, *FVR*, and all possible interactions between them were treated as predictor variables while *participant* was treated as a random ef-

fect.¹ Although normalized COG values have been used in previous studies to account for variation in vocal tract size (Dalola & Bridwell 2019, 2020), this proved to be redundant when *participant* was included as a random effect, and therefore, will not be used here. The proportion of the variance in the data explained by the random effects is reported in terms of the intraclass correlation coefficient (ICC), as calculated by the function *icc()* in the package *performance* (Lüdtke et al. 2021). Since post-hoc tests are not traditionally performed on mixed models (Levshina 2015), results below are interpreted in terms of 95% confidence intervals, as generated by the *effects* package (Fox 2003, Fox & Weisberg 2018, 2019). Visualizations were generated using *ggplot2* (Wickham 2016).

3 Results

3.1 COG

As presented in Table 2, the results of the COG model indicated a significant three-way interaction between *vowel*, *speaker group*, and *FVR*. While the complexity of this interaction makes interpreting the statistical results difficult, a visual presentation of the effects generated by this model (Figure 3) makes the patterns in the data clear. These are interpreted below, with all differences reported as statistically significant, reflecting results from 95% confidence intervals. In terms of variation present in the data, the ICC measure of 0.338 suggests relatively low levels of similarity between measurements taken from the same speaker, indicative of a high level of variability throughout.

Overall, it can be seen that COG increases as FVR increases, a pattern which is true of all vowels and both speaker groups, although the rate of increase differs across vowels. At low FVRs (i.e. 0%), /y,u/ > /i/ for L1 speakers ($p < .05$); L2 speakers exhibit no significant differences between vowels here, although /y/ trends higher than /i/ ($p = .071$). At high FVRs (i.e. 100%), /i/ > /y/ > /u/ for L1 speakers ($p < .05$); for L2 speakers, /i/ > /y,u/ ($p < .05$), although the difference between /y/ and /u/ is nearly significant ($p = .055$). At no point do L1 speakers differ from L2 speakers for a given vowel and FVR in terms of statistical significance.

¹Because repeated measures were taken for each word, initial models also included *word* as a random effect. However, examination of the ICCs for each model indicated that the grouping factor of *word* accounted for a very low amount of the total variation (1–3%). Since random effects with ICCs below 0.1 do not appreciably change the model (Vajargah & Nikbakht 2015), our models were therefore simplified to only include *participant* as a random effect.

Table 2: Mixed-effects linear regression model for COG

	Estimate	(Lower, upper 95% CI)	p-values	
(Intercept)	-301.75	(-535.03, -64.30)	.0124	*
Vowel /y/	944.41	(711.35, 1177.46)	.0000	***
Vowel /u/	936.01	(685.94, 1186.07)	.0000	***
FVR	4331.57	(4112.90, 4550.25)	.0000	***
Speaker group L2 French	272.51	(-148.23, 686.20)	.1970	
Vowel /y/*FVR	-2774.62	(-3111.76, -2437.46)	.0000	***
Vowel /u/*FVR	-3735.46	(-4116.90, -3354.01)	.0000	***
Vowel /y/*Speaker group L2 French	-186.28	(-628.13, 255.57)	.4086	
Vowel /u/*Speaker group L2 French	-742.73	(-1192.21, -293.26)	.0012	**
FVR*Speaker group L2 French	-280.07	(-715.64, 155.50)	.2075	
Vowel /y/*L2 French*FVR	76.08	(-564.30, 716.45)	.8159	
Vowel /u/*L2 French*FVR	840.39	(126.07, 1554.71)	.0211	*
Participant ICC	0.338			

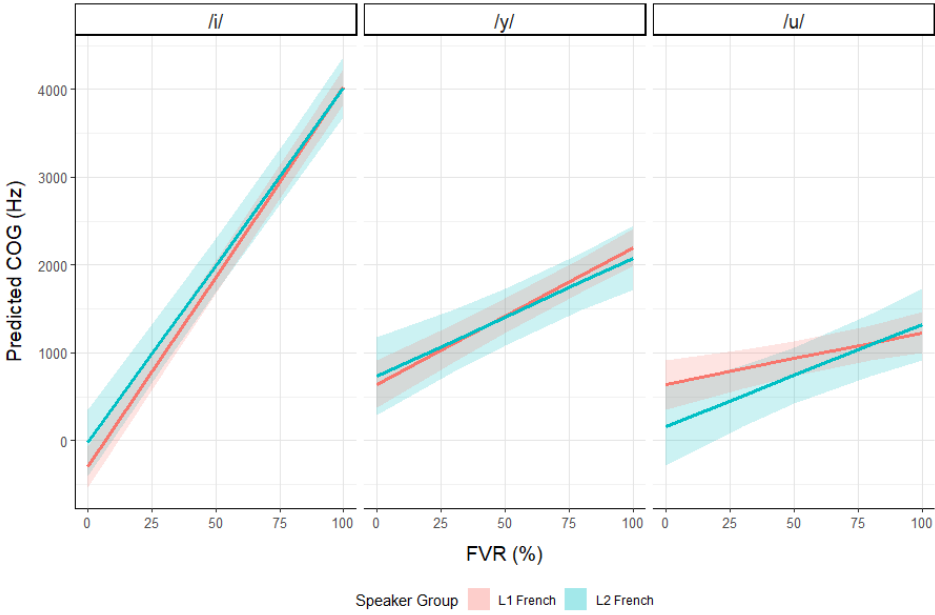


Figure 3: COG regression model effects

3.2 Standard deviation

As shown in Table 3, significant two-way interactions for *standard deviation* were present between *vowel* and *FVR* and *FVR* and *speaker group* while the interaction between *vowel* and *speaker group* was nearly significant. As visualized in Figure 4, patterns of standard deviation differed across vowels: it significantly increased with FVR for /i,u/ ($p < .05$), but decreased for /y/ ($p < .05$). As a result, /i,y/ > /u/ at low FVRs ($p < .05$), but at high FVRs /i/ > /u/ > /y/ ($p < .05$). In terms of speaker group, standard deviation increased with FVR for L1 speakers ($p < .05$) but not for L2 speakers ($p > .10$). At low FVRs, L2 speakers tended toward higher standard deviations than L1 speakers ($p = .096$); at high FVRs L1 speakers tended to show higher values ($p = .080$).

The ICC measure of 0.324 suggests relatively low levels of similarity between measurements taken from the same speaker, indicative of a high level of variability throughout.

Table 3: Mixed-effects linear regression model for standard deviation

	Estimate	(Lower, upper 95% CI)	p-values	
(Intercept)	1621.70	(1497.28, 1745.97)	.0000	***
Vowel /y/	-2.15.36	(-341.10, -89.62)	.0008	***
Vowel /u/	-510.55	(-645.48, -375.63)	.0000	***
FVR	480.69	(362.78, 598.61)	.0000	***
Speaker group L2 French	310.68	(89.88, 532.70)	.0061	**
Vowel /y/*FVR	-584.22	(-766.12, -402.30)	.0000	***
Vowel /u/*FVR	-19.64	(-225.45, 186.18)	.8517	
Vowel /y/*Speaker group L2 French	-20.74	(-259.16, 217.67)	.8646	
Vowel /u/*Speaker group L2 French	-226.14	(-468.67, 16.39)	.0676	.
FVR*Speaker group L2 French	-576.81	(-811.74, -341.86)	.0000	***
Vowel /y/*L2 French*FVR	14.93	(-330.60, 360.47)	.9325	
Vowel /u/*L2 French*FVR	297.88	(-87.56, 683.32)	.1298	
Participant ICC	0.324			

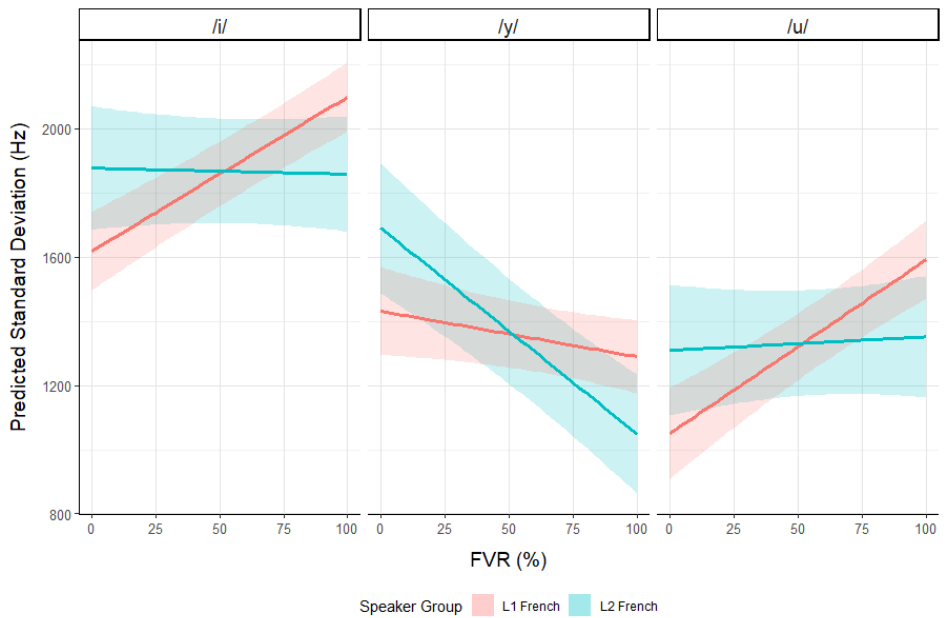


Figure 4: Standard deviation regression model effects

3.3 Skewness

As shown in Table 4, a significant interaction for *skewness* was present between *vowel*, *FVR*, and *speaker group*, an interaction which primarily centers around the results for /y/. As can be seen in Figure 5, there is a significant decrease in skewness as FVR increases ($p<.05$), except for L2 speakers' production of /y/ ($p>.10$), which stays relatively stable across FVRs. Correspondingly, predicted values for /u/ are greater than those for /i/ across FVRs ($p<.05$), but /y/ behaves differently, patterning with /i/ at low FVRs but with /u/ at high FVRs. Finally, L1 speakers show higher skewness values than L2 speakers at low FVRs for /y/ ($p<.05$). A tendency is also present for L1 speakers' skewness to be higher than that of L2 speakers at low FVRs for /i/ ($p=.060$).

Table 4: Mixed-effects linear regression model for skewness

	Estimate	(Lower, upper 95% CI)	<i>p</i> -values	
(Intercept)	9.254	(8.330, 10.170)	.0000	***
Vowel /y/	-0.450	(-1.410, 0.509)	.3575	
Vowel /u/	3.106	(2.076, 4.135)	.0000	***
FVR	-8.319	(-9.219, -7.419)	.0000	***
Speaker group L2 French	-2.315	(-3.970, -0.650)	.0064	**
Vowel /y/*FVR	3.392	(2.003, 4.780)	.0000	***
Vowel /u/*FVR	1.360	(-0.211, 2.931)	.0898	.
Vowel /y/*Speaker group L2 French	-0.618	(-2.438, 1.201)	.5054	
Vowel /u/*Speaker group L2 French	2.797	(0.946, 4.648)	.0031	**
FVR*Speaker group L2 French	1.582	(-0.210, 3.373)	.0835	.
Vowel /y/*L2 French*FVR	3.249	(0.612, 5.886)	.0158	*
Vowel /u/*L2 French*FVR	-2.250	(-5.191, 0.692)	.1339	
Participant ICC	0.299			

The ICC measure of 0.324 suggests relatively low levels of similarity between measurements taken from the same speaker, indicative of a high level of variability throughout.

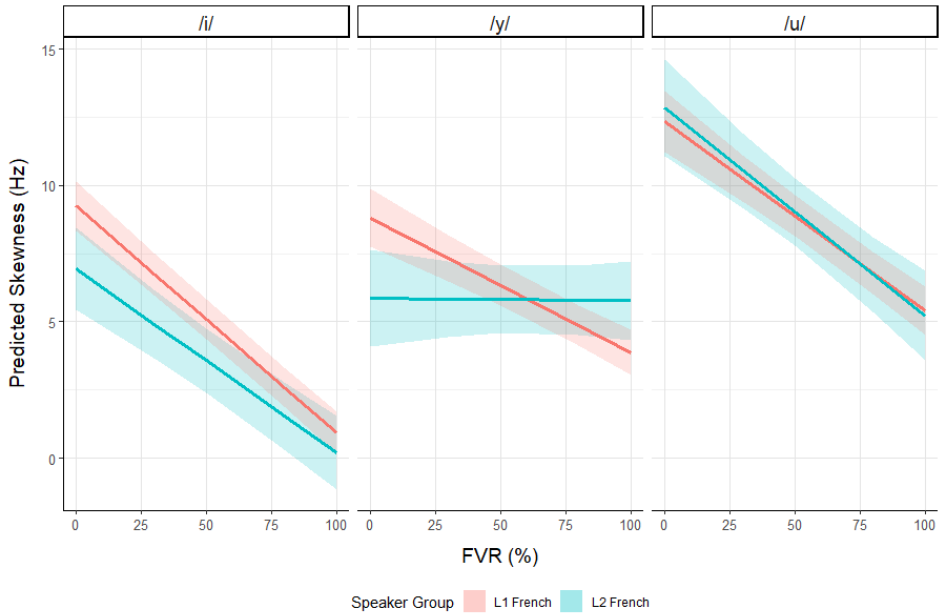


Figure 5: Skewness regression model effects

3.4 Kurtosis

As shown in Table 5, a trending three-way interaction for *kurtosis* was present between *vowel*, *FVR*, and *speaker group* (which later proved to be fully significant when /y/ was compared with /u/); a significant two-way interaction was also present between *vowel* and *speaker group*, and a trending interaction between *FVR* and *speaker group*.

As can be seen in Figure 6, patterns for kurtosis closely paralleled those for skewness; indeed, the two variables were very highly correlated ($r=0.911$). For L1 speakers, kurtosis significantly decreased with FVR for all vowels ($p<.05$); for L2 speakers, this pattern was only significant for /u/ ($p<.05$). At low FVRs, /u/ was significantly higher than /i,y/ for both speaker groups ($p<.05$); for high FVRs, /u/ > /i/ for L1 speakers ($p<.05$), while /y/ > /i/ for L2 speakers ($p<.05$). This reflects the fact that L2 /y/ exhibited a non-significant increase in kurtosis as FVR increased, a pattern in the opposite direction of other vowels. Finally, L1 speakers tended to produce PFFE with higher kurtosis for /i/ and /y/ at low FVRs ($p=.072, .080$).

The ICC measure of 0.266 suggests relatively low levels of similarity between measurements taken from the same speaker, indicative of a high level of variability throughout.

Table 5: Mixed-effects linear regression model for kurtosis

	Estimate	(Lower, upper 95% CI)	p-values	
(Intercept)	145.27	(115.50, 174.86)	.0000	***
Vowel /y/	-4.55	(-36.87, 27.77)	.7824	
Vowel /u/	81.65	(46.97, 116.33)	.0000	***
FVR	-130.58	(-160.88, -100.29)	.0000	***
Speaker group L2 French	-71.80	(-126.12, -17.30)	.0099	**
Vowel /y/*FVR	32.93	(-13.83, 79.68)	.1675	
Vowel /u/*FVR	-24.96	(-77.86, 27.94)	.3551	
Vowel /y/*Speaker group L2 French	-11.11	(-72.37, 50.15)	.7222	
Vowel /u/*Speaker group L2 French	74.35	(12.02, 136.68)	.0194	*
FVR*Speaker group L2 French	56.75	(-3.50, 117.00)	.0649	.
Vowel /y/*L2 French*FVR	80.08	(-8.72, 168.87)	.0772	.
Vowel /u/*L2 French*FVR	-82.76	(-181.54, 16.59)	.1027	
Participant ICC	0.266			

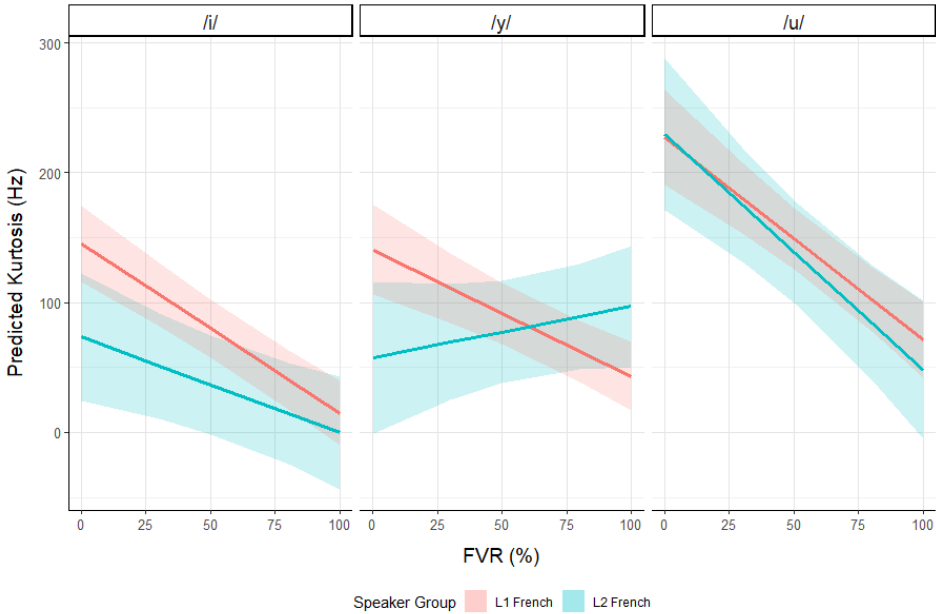


Figure 6: Kurtosis regression model effects

3.5 Intensity

Table 6: Mixed-effects linear regression model for intensity

	Estimate	(Lower, upper 95% CI)	<i>p</i> -values	
(Intercept)	40.847	(39.056, 42.645)	.0000	***
Vowel /y/	0.170	(-1.079, 1.419)	.7900	
Vowel /u/	1.266	(-0.074, 2.607)	.0641	.
FVR	8.073	(6.899, 9.246)	.0000	***
Speaker group L2 French	-3.974	(-6.546, -1.412)	.0024	**
Vowel /y/*FVR	3.617	(1.810, 5.424)	.0001	***
Vowel /u/*FVR	-1.292	(-3.337, 0.752)	.2153	
Vowel /y/*Speaker group L2 French	0.792	(-1.578, 3.162)	.5125	
Vowel /u/*Speaker group L2 French	4.267	(1.857, 6.676)	.0005	***
FVR*Speaker group L2 French	6.129	(3.783, 8.475)	.0000	***
Vowel /y/*L2 French*FVR	-2.087	(-5.521, 1.347)	.2335	
Vowel /u/*L2 French*FVR	-9.263	(-13.092, -5.434)	.0000	***
Participant ICC	0.592			

As shown in Table 6, a significant three-way interaction for *intensity* was present between *vowel*, *FVR*, and *speaker group*. The ICC for the intensity model, 0.592, was higher than for the four spectral moments, with inter-speaker differences accounting for more than half of the variation in the data. This reflects the fact that baseline loudness can markedly differ between speakers due to variation in microphone positioning and speaking style.

As shown in Figure 7, the most pronounced pattern was that intensity increased with FVR, a pattern which was significant for all vowels except L2 /u/ ($p < .05$). At low FVRs, L1 speakers exhibited no significant differences between vowels; L2 speakers produced /u/ with significantly higher intensities than /i/ ($p < .05$). At high FVRs, L1 speakers produced /y/ with higher intensity than /i,u/ ($p < .05$); L2 speakers produced /y/ higher than /u/ ($p < .05$) and tended to produce /i/ with higher intensity than /u/ ($p = .058$). There was also a tendency for L1 speakers to produce /i/ at low FVRs with greater intensity than L2 speakers ($p = .073$).

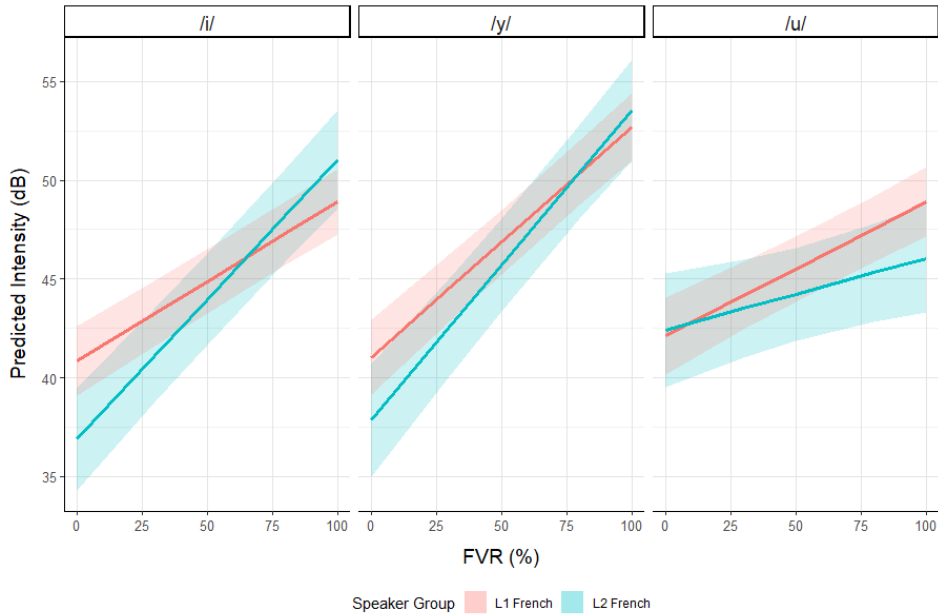


Figure 7: Intensity regression model effects

4 Discussion

4.1 Speaker Group Differences

Of the five parameters examined, categorical effects for speaker group were found only in the category of skewness and only for the vowel /y/, where L1 speakers exhibited significantly higher measures than L2s. Given that skewness is a metric of spectral tilt describing the distribution of frequencies occurring above or below the spectral mean, the higher measures reported for the L1 group indicate an L1 PFFE production for /y/ that is negatively skewed, because it contains a higher proportion of frequencies higher than the mean. Although skewness had not been examined in CF PFFE prior to this study, this finding parallels one reported by Munson et al. (2006) who found negatively skewed [s] to be a marker of American anglophone gay men's speech. The preponderance of higher frequencies in L1 speakers' negatively skewed /y/ may contribute to its increased saliency, rendering the phenomenon more easily detectable in L1 /y/ than in other L1 vowels and in all vowels where PFFE is produced by L2s. Future studies should seek to determine whether skewness plays a significant role in the perception of PFFE.

That no other significant main effects were found for speaker group in the remaining parameters is meaningful because it highlights the nuanced interaction of vowel type and FVR in conditioning PFFE among CF speakers and demonstrates the extent to which L2 PFFE is similar in nearly every perceptible phonetic way to L1 PFFE. It also marks a break with a previous study (Dalola & Bridwell 2020) that reported speaker group differences in the category of COG. A comparison of methodologies reveals that such differences likely stem from the treatment of speaker as a random effect in the current COG analysis in lieu of a simple normalization of COG scores. Future work should be explicitly designed to test these methods against one another to determine which holds not only more phonetic validity, but more sociophonetic validity.

4.2 Other Differences: The case of /y/

Previous research (Dalola & Bridwell 2020) reported divergent behavior for the treatment of /y/ across speakers groups. This analysis reports categorical speaker group differences for /y/ in terms of skewness (as discussed above), as well as for kurtosis and for intensity at higher FVRs. Given that kurtosis is a metric of the distribution of frequencies at the tails of distribution, the higher measures reported for the L2 group at higher FVRs indicate an L2 PFFE production for /y/ at higher FVRs that is less negatively kurtotic (the kurtosis scores are still negative) because it is characterized by a greater number of frequencies concentrated around the distribution mean than is found in L1 speakers, in different vowels and at lower FVRs. Although both L1 and L2 groups yielded negative kurtosis scores (meaning an overall greater number of frequencies were located towards the tails than towards the mean), the lesser degree of eccentricity in the frequencies of the L2 speakers suggests a higher degree of uniformity in L2 /y/ PFFE production than in L1.

/y/ also proved exceptional in terms of measures of standard deviation, where it was observed that both speaker groups exhibited a decrease from lower to higher FVRs, indicating a greater degree of variation in the frequencies at lower FVRs and a lesser degree of variation in the frequencies among higher FVRs. This finding contrasted with the other target vowels, which instead witnessed an increase in measures of standard deviation from lower to higher FVRs, indicating a greater degree of variation in the frequencies as FVRs increased.

PFFE for /y/ also exhibited a special status in terms of intensity at higher FVRs, with L1 speakers realizing it significantly louder than both /i/ and /u/, and L2 speakers realizing it at the same level of loudness as /i/, both of which were significantly louder than /u/. Taken together, this suggests that the higher overall

intensity of /y/ may be suggestive of a sort of “hyperspeech” (Lindblom 1990), in which CF speakers signal to their interlocutors their awareness of PFFE as a sociophonetic marker of polished and affected French, and, in doing so, overemphasize certain phonetic features at the expense of maximum articulatory effort. Dalola & Bridwell (2020) originally suggested that this behavior might be unique to L2 speakers, given the increased markedness of /y/ for non-native CF speakers and this group’s elevated measures of COG at lower FVRs (not corroborated here), but the present analysis indicates that the behavior may instead be characteristic of the vowel /y/ and shared across speaker groups. This theory is supported by the previous work examining L1 and L2 speakers’ perceptions of the variable, in which it was found that both L1 and L2 speakers construe PFFE as being associated with features of *TRUSTWORTHINESS* and *FORMALITY* (Dalola 2016), the second of which has notable social capital for CF speakers in their daily and/or professional life (the context in which they were sampled). Similar sociophonetic behavior has been found in white Southern Americans using hyperarticulated [hw] to index educatedness (Bridwell 2019), a phonetic behavior characterized by increased duration of the fricative portion of the segment.

4.3 Implications

In light of this study’s findings, we are now able to suggest four additional parameters to the definition of “sociophonetic competence” as laid out by Dalola & Bullock (2017) and as revised via the addition of COG measures by Dalola & Bridwell (2019, 2020). Previous work on the sociophonetic variable of PFFE in CF has demonstrated that it is not merely sufficient for L2 speakers to have awareness of a sociophonetic variable in their L2 to use it at similar rates or durations as their L1 counterparts, or even in the same types of pragmatic and phonological contexts. This study has instead identified an additional dimension of L2 mastery, namely that of phonetic quality of use, and defined it using four novel acoustic metrics relevant to the PFFE variable (SD, skewness, kurtosis, intensity – COG was first added in Dalola & Bridwell 2019). Such a mastery at the level of production would also imply a heightened sensitivity to the perception of these sound variations, affording speakers the ability to decode an additional layer of meaning in an L2, although it is likely that the perception of these phonetic differences precedes their target-like production (Flege & Hillenbrand 1984). The previous sociophonetic work that has examined various spectral moments and found them to work alone or in concert with other parameters to index information about the speaker – COG and *f*₀ as markers of masculine voices (Zimman 2017), negatively skewed /s/ (i.e. fronter, higher frequency) as a marker of gay

men's speech (Munson et al. 2006), intensity (loudness) values in PFFE segments as markers of negative affect (Dalola 2016) – suggests that the present findings may not only be indicative of speaker group status, but also constructs of gender, sexuality, and affect. Whereas the acoustic energy of PFFE realizations seems to vary allophonically for L1 French speakers, as predicted purely by phonological context and ease of articulation, it seems to vary sociophonetically and pragmatically for L2 speakers, as conditioned by the desire for speakers to signal their sociophonetic awareness to native listeners at structurally and pragmatically acceptable moments.

5 Conclusion

Future studies will sample advanced L2 French populations more robustly and subdivide their level of advancedness via quantitative measures, i.e., the Bilingual Language Profile (Birdsong et al. 2012). The high levels of variability in L2 PFFE documented here may be optimized by a more rigorous and nuanced classification of speaker group. Additionally, we propose that the current findings be tested via a series of perceptual studies that investigate the pragmatic values of PFFE with differing and controlled measures of COG, SD, skewness, kurtosis, and intensity in both L1- and L2-French populations. In that way, we will be able to isolate which phonetic components of PFFE contribute most reliably and meaningfully to perceptual differences and which ones represent mere physiological variation. Finally, it is imperative that the phonetic and lab phonological approaches currently popular in this type of research seek to combine with or add methods from third-wave sociolinguistics (Eckert 2012), so that the constructs of gender, sexuality, and affect may be investigated more closely with relation to the PFFE variable.

Acknowledgements

We would like to thank our reviewers and the attendees at LSRL50 for their detailed and thoughtful feedback. Any remaining errors are our own.

References

- Bates, Douglas, Martin Mächler, Ben Bolker & Steve Walker. 2015. Fitting linear mixed-effects models using lme4. *Journal of Statistical Software* 67. 1–48.

- Birdsong, David, Libby M. Gertken & Mark Amengual. 2012. *Bilingual language profile: An easy-to-use instrument to assess bilingualism*. COERLL, University of Texas at Austin. <https://sites.la.utexas.edu/bilingual/> (31 August, 2016).
- Boersma, Paul & David Weenink. 2018. *Praat: Doing phonetics by computer*. [Computer program]. Version 6.0.37.
- Bridwell, Keiko. 2019. *The distribution of [ɹ]: An acoustic analysis of sociophonetic factors governing the wine-whine merger in Southern American English*. Columbia, SC: University of South Carolina. (MA thesis).
- Candea, Maria. 2012. Au journal de RFI-chhh et dans d'autres émissions radiodiffusée-chhhs: Les épithèses consonantiques fricatives. *Le discours et la langue* 2(2). 136–149.
- Candea, Maria, Martine Adda-Decker & Lori Lamel. 2013. Recent evolution of non-standard consonantal variants in French broadcast news. In Frédéric Bimbot, Cécile Fougeron & François Pellegrino (eds.), *Interspeech*, 412–416. Lyon: ISCA.
- Dalola, Amanda. 2014. *Un drôle de bruit_hhh: A sociophonetic examination of the production and perception of final vowel devoicing among L1 and L2 speakers of French*. Austin, TX: University of Texas at Austin. (Doctoral dissertation). <http://hdl.handle.net/2152/31395>.
- Dalola, Amanda. 2015a. *An acoustic profile of phrase-final vowel devoicing in Continental French*. LSRL45, Campinas, Brazil [Oral Presentation].
- Dalola, Amanda. 2015b. The role of vowel type, preceding consonant and lexical frequency on final vowel devoicing in Continental French. In Sebastian Möller & Hermann Ney (eds.), *Proceedings of the 18th International Congress of Phonetic Sciences*. Glasgow: ISCA. <https://www.internationalphoneticassociation.org/icphs-proceedings/ICPhS2015/Papers/ICPHS0649.pdf>.
- Dalola, Amanda. 2016. *Bossy is in the ear of the beholder: When L1 and L2 French speakers perceive final vowel devoicing*. NWAV45, Vancouver [Oral Presentation].
- Dalola, Amanda & Keiko Bridwell. 2019. The shape of [u]: Towards a typology of final vowel devoicing in Continental French. In Sasha Calhoun, Paola Escudero, Marija Tabain & Paul Warren (eds.), *Proceedings of the 19th International Congress of Phonetic Sciences*. Canberra: Australasian Speech Science & Technology Association. https://www.internationalphoneticassociation.org/icphs-proceedings/ICPhS2019/papers/ICPhS_1223.pdf.
- Dalola, Amanda & Keiko Bridwell. 2020. Redefining sociophonetic competence: Mapping COG differences in phrase-final fricative epithesis in L1 and L2 speakers of French. *Languages* 5(4). 59.

- Dalola, Amanda & Barbara E. Bullock. 2017. On sociophonetic competence. *Studies in Second Language Acquisition* 39(4). 769–799.
- Eckert, Penelope. 2012. Three waves of variation study: The emergence of meaning in the study of sociolinguistic variation. *Annual review of Anthropology* 41. 87–100.
- Fagyal, Zsuzsanna. 2010. *Accents de banlieue: Aspects prosodiques du français populaire en contact avec les langues de l'immigration*. Paris: L'Harmattan.
- Fagyal, Zsuzsanna & Christine Moisset. 1999. Sound change and articulatory release: Where and why are high vowels devoiced in Parisian French. In Géza Gordos (ed.), *Proceedings of the 14th International Congress of Phonetic Sciences*, 309–312. San Francisco, CA: ISCA.
- Flege, James E. 1987. The production of “new” and “similar” phones in a foreign language: Evidence for the effect of equivalence classification. *Journal of phonetics* 15(1). 47–65.
- Flege, James E. 1995. Second language speech learning: Theory, findings, and problems. In Strange Winifred (ed.), *Speech perception and linguistic experience: Issues in cross-language research*, 233–277. Baltimore: York Press.
- Flege, James E. & James Hillenbrand. 1984. Limits on phonetic accuracy in foreign language speech production. *The Journal of the Acoustical Society of America* 76(3). 708–721.
- Fónagy, Ivan. 1989. Le français change de visage? *Revue Romane* 24(2). 225–254.
- Forrest, Karen, Gary Weismer, Paul Milenkovic & Ronald N. Dougall. 1988. Statistical analysis of word-initial voiceless obstruents: Preliminary data. *The Journal of the Acoustical Society of America* 84(1). 115–123.
- Fox, John. 2003. Effect displays in R for generalised linear models. *Journal of statistical software* 8(15). 1–27.
- Fox, John & Sanford Weisberg. 2018. Visualizing fit and lack of fit in complex regression models with predictor effect plots and partial residuals. *Journal of Statistical Software* 87(1). 1–27.
- Fox, John & Sanford Weisberg. 2019. *An R companion to applied regression*. Thousand Oaks: Sage publications.
- Kuznetsova, Alexandra, Per B. Brockhoff & Rune H. B. Christensen. 2017. Lmertest package: Tests in linear mixed effects models. *Journal of Statistical Software* 82(13). 1–26. DOI: 10.18637/jss.v082.i13.
- Levshina, Natalia. 2015. *How to do linguistics with R: Data exploration and statistical analysis*. Amsterdam: John Benjamins.
- Lindblom, Björn. 1990. Explaining phonetic variation: A sketch of the H&H theory. In William J Hardcastle & Alain Marchal (eds.), *Speech production and speech modelling*, 403–439. Berlin: Springer.

- Lüdecke, Daniel, Mattan S. Ben-Shachar, Indrajeet Patil, Philip Waggoner & Dominique Makowski. 2021. Performance: An R package for assessment, comparison and testing of statistical models. *Journal of Open Source Software* 6(60). 3139. DOI: 10.21105/joss.03139.
- Martin, Pierre. 2004. Dévoisement vocalique en français. *La linguistique* 40(2). 3–22.
- Munson, Benjamin, Elizabeth C. McDonald, Nancy L. DeBoe & Aubrey R. White. 2006. The acoustic and perceptual bases of judgments of women and men's sexual orientation from read speech. *Journal of Phonetics* 34(2). 202–240.
- Osborne, Jason W. & Anna B. Costello. 2009. Best practices in exploratory factor analysis: Four recommendations for getting the most from your analysis. *Pan-Pacific Management Review* 12(2). 131–146.
- Paternostro, Roberto. 2008. Le dévoisement des voyelles finales. *Rassegna italiana di Linguistica applicata* 3(40). 129–158.
- R Core Team. 2018. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing. Vienna. <https://www.R-project.org/>.
- Smith, Caroline L. 2002. Prosodic finality and sentence type in French. *Language and Speech* 45(2). 141–178.
- Smith, Caroline L. 2003. Vowel devoicing in contemporary French. *Journal of French Language Studies* 13(2). 177–194.
- Smith, Caroline L. 2006. *In and out of the laboratory: Using a variety of data sources to study variation source in speech*. online. Poster. Barcelona. <https://www.unm.edu/~caroline/CSmithLabPhon06.pdf>.
- Toda, Martine. 2007. Speaker normalization of fricative noise: Considerations on language-specific contrast. In Jürgen Trouvain (ed.), *16th international congress on phonetic sciences*, 825–828. Saarbrücken: Univ. des Saarlandes.
- Toda, Martine. 2009. *Etude articulatoire et acoustique des fricatives sibilantes*. Paris: Université Sorbonne Nouvelle – Paris 3. (Doctoral dissertation).
- Toda, Martine, Shinji Maeda & Kiyoshi Honda. 2010. Formant-cavity affiliation in sibilant fricatives. In Susanne Fuchs (ed.), *Turbulent sounds: An interdisciplinary guide*, 343–374. Berlin: Walter de Gruyter.
- Vajargah, Kianoush Fathi & Masoom Nikbakht. 2015. Application REML model and determining cut off of ICC by multi-level model based on Markov chains simulation in health. *Indian Journal of Fundamental and Applied Life Sciences* 5. 2231–6345.
- Wickham, Hadley. 2016. *Ggplot2: Elegant graphics for data analysis*. 2nd edn. (Use R!). New York: Springer. DOI: 10.1007/978-3-319-24277-4. <http://ggplot2.org>.

Zimman, Lal. 2017. Gender as stylistic bricolage: Transmasculine voices and the relationship between fundamental frequency and /s/. *Language in Society* 46(3). 339–370.