Article

Perceptual and Motor Skills
2021, Vol. 128(1) 153–177

DOI: 10.1177/0031512520973514
journals.sagepub.com/home/pms
SAGE

Cross-Gender Differences in English/French Bilingual Speakers: A Multiparametric Study

Erwan Pépiot[1] and Aron Arnold[2]

Abstract

The present study concerns speech productions of female and male English/French bilingual speakers in both reading and semi-spontaneous speech tasks. We investigated various acoustic parameters: average fundamental sound frequency (F0), F0 range, F0 variance (*SD*), vowel formants (F1, F2, and F3), voice onset time (VOT) and H1-H2 (intensity difference between the first and the second harmonic frequencies, used to measure phonation type) in both languages. Our results revealed a significant effect of gender and language on all parameters. Overall, average F0 was higher in French while F0 modulation was stronger in English. Regardless of language, female speakers exhibited higher F0 than male speakers. Moreover, the higher average F0 in French was larger in female speakers. On the other hand, the smaller F0 modulation in French was stronger in male speakers. The analysis of vowel formants showed that overall, female speakers exhibited higher values than males. However, we found a significant cross-gender difference on F2 of the back vowel [u:] in English, but not on the vowel [u] in French. VOT of voiceless stops was longer in Female speakers in both languages, with a greater difference in English. VOT contrast between voiceless stops and their voiced counterparts was also significantly longer in female speakers in both languages. The scope of this cross-gender difference was greater in English. H1-H2 was higher in female speakers in both languages, indicating a breathier phonation type. Furthermore, female speakers tended to exhibit smaller H1-H2 in

[1]TransCrit, University of Paris 8, Saint-Denis, France
[2]VALIBEL, Catholic University of Louvain, Louvain-la-Neuve, Belgium

Corresponding Author:
Erwan Pépiot, Transcrit, University of Paris 8, 2 rue de la liberté, Saint-Denis 93526, France.
Email: erwan.pepiot@free.fr

French, while the opposite was true in males. This resulted in a smaller cross-gender difference in French for this parameter. All these data support the idea of language- and gender-specific vocal norms, to which bilingual speakers seem to adapt. This constitutes a further argument to give social factors, such as gender dynamics, more consideration in phonetic studies.

Keywords

speech production, speech and gender, voice and gender, cross-language variation, bilingualism

Introduction

Over recent decades, many studies have addressed acoustic differences between female and male speech. Most studies focused on fundamental frequency (F0) and resonant frequencies (especially vowel formants). These two parameters are considered the main cross-gender differences in speech production. Some authors investigated cross-gender F0 differences in several different languages (e.g. Traunmüller & Eriksson, 1995), but very few researchers have been interested in intra-individual variations that occur when bilingual speakers switch from one language to another. We suggest that the study of intra-individual variations in this circumstance will further challenge the restrictive perspective that F0 and vowel formant frequencies are mostly dependent on the speaker's anatomy and will support, instead, a more dynamic understanding that considers the importance of culture-related gender differences.

Acoustically, F0 is usually lower in males' voices (Boë et al., 1975) while vowel formants are typically at higher frequencies in female's voices. These acoustic gender differences are partly due to developmental differences in the human vocal apparatus that emerge during puberty. At that developmental period, males' vocal folds are becoming longer and thicker, and the vocal tract becomes longer into adulthood (Abitbol et al., 1999; Kahane, 1978). This is one of the reasons why vocal folds typically vibrate more slowly in male than in female speakers, and why resonant frequencies tend to be higher in females than males. Other factors such as age (Honjo & Isshiki, 1980) and cigarette consumption (Matar et al., 2016) have a further effect on the vocal folds, modifying the average F0.

But anatomy and physiological factors cannot account for all of the differences between female and male voices (Arnold, 2015; Pépiot, 2014a). Each speaker has a unique vocal apparatus with a given shape (influencing both F0 and resonant frequencies), and the individual uses this vocal apparatus to index a specific gender identity. Therefore, a voice is never solely a reflection of one's anatomy, but it is also the result of a culturally gendered performance. For instance, vowel formant frequencies can be modified by adjusting the position of the tongue and the protrusion of the lips. Typically, women use certain articulatory practices to produce relatively clear and high-pitched voices, while men use other practices to achieve relatively dark and low-pitched voices (Arnold, 2016). Such phenomena can be viewed as speech-related gender norms.

It is also well known that cross-gender differences on F0 and resonant frequencies are language dependent (Johnson, 2005; Van Bezooijen, 1995). In Wu language (a Shanghai dialect), for example, gender differences in average F0 appear to be unusually small (Rose, 1991). As to vowel formants, a contrastive study showed that the scope of cross-gender differences is not the same in French and in American English, in that there is a particularly small difference on back vowels in French (Pépiot, 2015).

Other acoustic parameters such as F0 range, or more generally, F0 modulation, have been described as loci for cross-gender differences. According to Austin (1965) and Lakoff (1975, p. 56), female speakers, compared to male speakers, tend to use greater F0 modulation and range. However, these results are still debated. On the one hand, when using the semitone scale, which is more representative of human sound perception than the Hertz scale, Henton (1989a, 1995) found no significant female/male differences in F0 modulations among American English speakers. On the other hand, Pépiot (2014b), using the same method, found significant cross-gender differences in French speakers in that female speakers actually modulated their speech more than males. Such findings suggest that cross-gender differences on F0 range and modulation may be related to the speaker's language.

Another parameter, phonation type, might also be related to the speaker's gender. Female voices have often been considered breathier (i.e. having a greater glottal open quotient –GOQ) than male voices (Henton & Bladon, 1985; Klatt & Klatt, 1990). Male voices, at least among American English speakers, are typically creakier (i.e. having a very low GOQ) than female voices (Henton, 1989b). However, recent studies have suggested that some young American female speakers now tend to use a creaky voice quality as well (Wolk et al., 2012). Moreover, these observations have varied slightly from one study to another, from one language to another (Pépiot, 2014b) and also depend on the acoustic parameters used to estimate phonation type. According to Gordon and Ladefoged (2001), intensity differences between H1 (first harmonic) and H2 (second harmonic) may be a reliable measure, if used properly (i.e., only in open vowels and without nasality in adjacent sounds - see Klatt & Klatt, 1990). Harmonic frequencies are integral multiples of the fundamental frequency. The lower harmonic is called H1, the next one H2, *etc*. Harmonics are visible on spectral slices, such as in Figure 2, and can be more or less intense.

Several authors have investigated cross-gender differences in voice onset time (VOT). Among English speakers, there is a general tendency for female speakers to produce longer VOTs than males, even though those values have varied noticeably from study to study. For instance, Swartz (1992) showed that VOTs were significantly shorter in English male speakers on alveolar plosives [t] and [d]. Similar results were found on voiceless plosives in English speakers (Robb et al., 2005; Whiteside & Irving, 1997). In other languages, results have not been that clear. Male Korean speakers have been found to use either longer or equivalent VOTs relative to female Korean speakers (Oh, 2011). Swedish speakers showed no significant cross-gender VOT differences (Karlsson et al., 2004). Among Parisian French speakers, Pépiot (2016) found that female speakers exhibited longer VOTs on voiceless stops and shorter VOTs (i.e. longer *pre-voicing*) on voiced stops than males.

What about bilingual speakers, then? Do bilingual speakers adapt to gender-related norms in different languages? These questions have not sparked much past interest and they have not been thoroughly investigated. In this paper, we define *bilingual speakers* as those who use two (or more) languages in their everyday lives (Grosjean & Li, 2013). A few studies conducted on bilingual speakers have shown that their average F0 depended on the language they used (Altenberg & Ferrand, 2006; Lee & Van Lancker Sidtis, 2017). Similar results were found for F0 range (Mennen et al., 2012). Altenberg and Ferrand (2006) showed that Russian L1/English L2 bilingual female speakers tended to exhibit a lower F0 in English. However, as their analysis included no males, we cannot discern whether this variation was an adaptation to gender norms or simply to language norms (unrelated to gender).

In the present study, we sought to investigate the productions of English L1/French L2 bilingual speakers of both genders, by measuring several acoustic parameters (F0, vowel formants, VOT, H1-H2) in two different speaking conditions: read speech and spontaneous speech. In phonetic sciences, the expression *read speech* refers to speech elicited from written material, whereas *spontaneous speech* refers to unprepared speech produced in interactive contexts, such as conversation, interview, etc. We hypothesized that bilingual speakers would adapt their vocal practices to the gender norms of the language they were using.

Method

Participants

Twelve English L1/French L2 bilingual speakers (6 women, 6 men) were recorded for this study. These participants were North-Eastern American speakers who had been living in Paris for several years. All of them used French daily and self-reported their French fluency level to be superior or equal to "3" on a scale from 0–5 (questionnaire inspired by Grosjean & Li, 2013). The two (female and male) groups can be considered homogenous since one man and one woman self-reported a "3" on this scale, while all other participants in both groups reported "4" or "5." Moreover, none of the speakers reported the regular use of any L3.

These participants were 29–54 years old ($SD = 7.6$ years) when the recording took place. The average age for male speakers was 40 ($SD = 8.2$) and, for female speakers, was 41.8 ($SD = 4.5$). All participants were non-smokers, and none reported any speech disorder. Each participant gave their written informed consent before engaging in the experiment, and all received a USB memory stick for their participation in the study. They were informed that all the data collected would be processed anonymously. The protocol, consisting of producing a simple and non-invasive audio recording, was approved by our laboratory's ethical committee.

Linguistic Material

This study was based on an analysis of French and English linguistic material, collected through three different tasks performed in both languages: (a) reading disyllabic (pseudo-)words, (b) reading sentences, and (c) a semi-spontaneous monologue.

For the dissyllabic reading condition, speakers had to read 12 sentences containing dissylabic words. We selected the dissyllabic (pseudo-)words based on two main criteria: (a) making the two corpora as similar as possible, and (b) limiting the number of combinations by choosing only the most relevant phonemes (i.e./i//a//u/vowels in French and/i:/,/æ/,/u:/vowels in English on the first syllable), while holding the last consonant/vowel (CV) sequence constant:/pi/ was chosen as this sound sequence can occur in word-final position in both languages. Thirty-three (C)VCV words were chosen for each language:

- /C (plosive) – V – p – i/combinations:/pipi/,/papi/,/pupi/,/bipi/,/bapi/,/bupi/, /tipi/,/tapi/,/tupi/,/dipi/,/dapi/,/dupi/,/kipi/,/kapi/,/kupi/,/gipi/,/gapi/,/gupi/ for the French corpus,/pi:pi/,/pæpi/,/pu:pi/,/bi:pi/,/bæpi/,/bu:pi/,/ˈti:pi/, /ˈtæpi/,/ˈtu:pi/,/ˈdi:pi/,/ˈdæpi/,/ˈdu:pi/,/ˈki:pi/,/ˈkæpi/,/ˈku:pi/,/ˈgi:pi/,/ˈgæpi/, /ˈgu:pi/for the English corpus.
- /C (fricative) – V – p – i/combinations:/sipi/,/sapi/,/supi/,/zipi/,/zapi/,/zupi/, /ʃipi/,/ʃapi/,/ʃupi/,/ʒipi/,/ʒapi/,/ʒupi/for the French corpus,/ˈsi:pi/,/ˈsæpi/, /ˈsu:pi/,/ˈzi:pi/,/ˈzæpi/,/ˈzu:pi/,/ˈʃi:pi/,/ˈʃæpi/,/ˈʃu:pi/,/ˈʒi:pi/,/ˈʒæpi/,/ˈʒu:pi/ for the English corpus.
- /V – p – i/combinations:/ipi/,/api/,/upi/for the French corpus,/ˈi:pi/,/ˈæpi/,/ˈu:pi/for the English corpus.

There is no phonological lexical stress in French (Hirst et al., 2001); but, within the frame sentence used for the recordings, speakers have naturally produced an emphatic stress on the first syllable of each experimental word.

Regarding the sentence reading task, each speaker also had to read 12 sentences in English (such as *"When the weather is cold and rainy, I'd rather stay at home."; "My sister told me she'd come by tomorrow."; "If you do that again, I'll call the police!"; . . .*) and 12 similar sentences in French *("Quand il fait froid et qu'il pleut, je préfère rester chez moi."*; *"Ma soeur m'a dit qu'elle allait passer demain."*; *"Si tu refais ça, j'appelle la police!"*; . . .)

Regarding the semi-spontaneous monologue, the speakers were invited to talk about their last vacation for 1–2 minutes, initiating this narration with the following English/French sentence: *"Tell me about your last vacation."* and "*Parlez-moi de vos dernières vacances."*

Recording Procedure

Recordings took place in anechoic chambers at the University of Paris 8 and the University of Paris 3, using a digital recorder *Edirol R09-HR* by Roland (frequency response range: 20 Hz – 40 kHz). Each participant was asked to perform the three speaking tasks as follows:

- Read dissyllabic words presented in an orthographical transcription on a sheet of paper and placed into a sentence frame in order to make prosodic parameters consistent: "He said '*WORD*' twice" for the English corpus and "Il a dit '*MOT*' deux fois" for the French one. Participants had to read each item twice.
- Read sentences (two readings per item).
- Give a semi-spontaneous narration about the participant's last vacation.

Participants performed these tasks in both French and English. Half of the speakers started with French, and the other half started with English, in order to neutralize possible biases caused by the order of speaking in the different languages (see Altenberg & Ferrand, 2006).

Data Analysis

We conducted the acoustic analysis in *Praat* (Boersma & Weenink, 2017) and analyzed the following speech parameters in the reading of sentences and the spontaneous speech tasks:

- Average F0 (average voice pitch)
- F0 range (i.e., the difference between the highest and lowest F0 within a given linguistic unit)
- F0 *SD* or F0 variance showing F0 modulation as the mean difference between each point of the F0 curve and the average F0

We obtained these data by creating a "pitch file" for each sentence/discourse and then collecting the values in the "pitch info" window. F0 range and F0 *SD* were measured in both Hertz and semitones. This scale was particularly appropriate because it takes into account the pitch variations as perceived by human listeners (Henton, 1995). For instance, an increase in frequency by 20 Hz between 100 and 120 Hz results in a larger melodic movement (+3.16 semitones) than a modulation between 200 Hz and 220 Hz (+1.65 semitones).

We performed additional acoustic measurements on dissyllabic words (33 individual words * 2 repetitions per speaker) as follows:

- F1, F2, F3 values (in Hz) on the first syllable vowels, collected manually, using spectrograms and spectra, in a central and stable portion of the vowels.
- VOT (in ms) of initial plosive consonants. This corresponded to the time frame between the release of the plosive consonant and the beginning of voicing. If voicing started before the release, as is typical in French voiced plosives, the value was negative. These data were measured manually on spectrograms.
- H1-H2 difference (in dB) in open vowels.

This last measurement gave an indication of phonation type. The relative strength of H1 is positively correlated with glottal open quotient (*GOQ*) (Klatt & Klatt, 1990). Nevertheless, H1-H2 can only be measured on open vowels; F1 would otherwise distort the results (Klatt & Klatt, 1990). Thus, vowel [a] for French and vowel [æ] for English were the only ones considered.

We made a five-period selection (i.e. five vibrations of the vocal folds) on a central portion of each open vowel. As shown in Figure 1, we displayed the

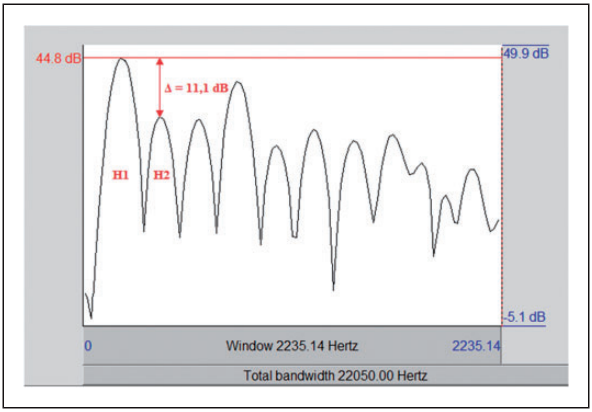

Figure 1. Measurement of H1-H2 Intensity Differences on Vowel [æ] Extracted From Word [æpi] Produced by a Female Speaker.

corresponding spectrum and we then manually calculated the difference between H1 and H2 intensity (in dB).

We subjected all these data to inferential statistical analysis, using two-way (two genders X two languages) analyses of variance (ANOVAs) in order to compare speech parameter differences between the two gender groups, the two languages spoken, and the interaction of these parameters. These tests were conducted in StatView 5.0. We used an alpha level of $p < .01$ to detect statistical significance. We also checked that the statistical power was above 0.8 for each performed analysis.

Results

Average F0

Analyses for average F0 for female and male speakers in read sentences are presented in Table 1, below.

As expected, average F0 was higher among female than male participants in both languages. The two-factor ANOVA ("*spoken language*" and "*gender*") showed a statistically significant effect of both *spoken language*, $F(1,572) = 25.566$, $p < .0001$, and *speaker's gender*, $F(1,572) = 2897.3$, $p < .0001$ on average

Table 1. Average F0 in Hertz (Hz) for Female (F) and Male (M) Speakers in Read Sentences (12 x 2 Occurrences) as a Function of the Spoken Language (English or French).

	Average F0 – Reading task (Hz)		
Speaker	English	French	% diff. FR/EN
F1	195	211	+8.28
F2	224	234	+4.29
F3	176	192	+8.68
F4	201	218	+8.37
F5	186	205	+10.20
F6	187	206	+10.01
F average	195	211	+8.17
M1	113	112	−1.29
M2	81	83	+2.63
M3	120	121	+1.11
M4	106	103	−3.49
M5	129	129	−0.10
M6	108	119	+9.77
M average	110	111	+1.33

The variation (in %) between English and French languages for each individual speaker is indicated in the right column.

F0. Moreover, there was a significant interaction between gender and language, $F(1{,}572) = 17.712$, $p < .0001$, meaning that female and male speakers did not adapt their average F0 the same way when they switched from one language to the other. There was an apparent cross-gender variation in the use of this acoustic parameter as a function of the spoken language, such that female speakers exhibited higher average F0 in French (M F0 = 211 Hz, $SD = 15.1$) than English (M F0 = 195 Hz, $SD = 12.9$), while male speakers exhibited similar average F0s. As explained earlier, our participants were also requested to produce semi-spontaneous speech. Average F0 values for these 1-2 minute speech sequences are presented in Table 2.

Overall, these results are consistent with what was found on read sentences results, with lower average F0 in male participants for both languages and an increase in females' average F0 in French when compared to English (+2.76%). The two-factor ANOVA ("*spoken language*" and "*gender*") for average F0 in semi-spontaneous speech showed significant main effects for both language, $F(1{,}476) = 7.059$, $p < .01$ and gender, $F(1{,}476) = 6062.193$, $p < .0001$. The interaction between language and gender approximated but did not reach statistical significance ($F(1{,}476) = 3.816$, $p = .0513$) as there was only a tendency for female speakers to show a relative increase in average F0 in French, compared to English.

Table 2. Average F0 of Female and Male Speakers on Semi-Spontaneous Speech, as a Function of the Spoken Language.

	Average F0 – Semi-spontaneous (Hz)		
Speaker	English	French	% diff. FR/EN
F1	179	189	+5.47
F2	190	195	+2.95
F3	167	175	+4.73
F4	193	197	+2.13
F5	173	177	+2.25
F6	184	182	−0.98
F average	181	186	+2.76
M1	104	105	+0.86
M2	74	73	−1.62
M3	103	105	+2.43
M4	99	99	+0.20
M5	121	121	+0.50
M6	99	100	+1.52
M average	100	101	+0.65

The variation (in %) between English and French languages for each individual speaker is indicated in the right column.

F0 Range and SD

F0 range (in Hertz and semitones) and its average standard deviation (in Hertz and semitones) on read sentences are shown in Table 3.

The F0 range tended to be reduced on read sentences when the bilingual participants spoke French versus English, as they averaged a reduction of 11.89% (in semitones) among males and 14.36% (in semitones) among females. Participants also tended to exhibit smaller *SDs* when speaking in French versus English, with this phenomenon greater among males, (*SD* in semitones decreased by 15.61%) than females (14.26%).

The two-factor ANOVA ("*spoken language*" and "*gender*") for read sentences conducted on F0 range values (in semitones) confirmed significant main effects of both language ($F(1,572) = 18.823$, $p < .0001$) and gender ($F(1,572) = 340.109$ with $p < .0001$). This pattern held for the two-factor ANOVA conducted on *SD* (in semitones), again with significant main effects for both *language* ($F(1,572) = 44.087$, $p < .0001$) and *gender* ($F(1,572) = 57.530$, $p < .0001$).

Regarding semi-spontaneous speech, the F0 range (in Hz and semitones) and F0 standard deviation (in Hz and semitones) are shown in Table 4. When one observes the *SD* differences between the two languages expressed in semitones

Table 3. Average Values of F0 Range in Hertz (Hz) and Semitones (st), Standard Deviation (SD) of F0 in Hertz and Semitones, for Both Female (F) and Male (M) Speakers on Read Sentences (12 x 2 Occurrences) as a Function of the Spoken Language (English or French).

	Read sentences - EN				Read sentences - FR				
Speaker	F0 ran. (Hz)	F0 ran. (st)	SD (Hz)	SD (st)	F0 ran. (Hz)	F0 ran. (dt)	SD (Hz)	SD (st)	% diff. FR/EN SD (st)
F1	218.28	20.37	50.56	4.85	219.36	19.06	40.53	3.60	**−25.76**
F2	233.25	20.21	46.37	3.79	203.33	15.92	39.46	2.97	**−21.76**
F3	166.34	16.71	32.56	3.22	165.56	14.95	25.24	2.29	**−28.90**
F4	201.91	17.93	38.39	3.23	224.59	19.56	41.46	3.54	**+9.60**
F5	182.75	16.61	38.00	3.45	162.96	14.00	33.85	2.88	**−16.54**
F6	173.64	16.88	28.12	2.68	213.91	18.91	33.74	2.92	**+8.88**
F aver.	**196.03**	**18.12**	**39.00**	**3.54**	**198.28**	**17.07**	**35.71**	**3.03**	**−14.26**
M1	101.92	15.30	23.61	3.51	88.56	13.96	19.00	3.00	**−14.59**
M2	54.12	10.68	9.78	2.01	53.50	10.47	10.53	2.14	**+6.27**
M3	91.55	13.82	22.03	3.24	82.75	11.49	19.19	2.62	**−19.05**
M4	79.84	12.50	20.13	3.16	72.43	11.86	16.91	2.79	**−11.55**
M5	94.08	11.95	24.41	3.13	78.71	10.32	18.48	2.41	**−23.25**
M6	76.66	12.98	16.85	2.73	69.23	9.94	14.87	2.06	**−24.84**
M aver.	**83.03**	**12.87**	**19.47**	**2.96**	**74.20**	**11.34**	**16.50**	**2.50**	**−15.61**

The variation of SD in semitones (expressed in %) between English and French languages for each individual speaker is indicated in the right column.

(last column in Table 4), one can see that overall, female and male participants modulated less when speaking in French than in English. But this tendency was more salient in male than in female speakers – the *SD* decrease from English to French was 15.85% in male speakers' speech, whereas it was only 3.41% among female speakers. This result was consistent with our results from the read speech task.

A two-factor ANOVA ("*spoken language*" and "*gender*") on the F0 *SD* parameter showed a significant effect of language, $F(1{,}476) = 29.353$, $p < .0001$. The same was true for gender, $F(1{,}476) = 11.371$, $p < .001$. Moreover, there was a significant interaction between the two factors ($F(1{,}476) = 14.097$, $p < .001$), indicating that female and male speakers did not change their modulations in the same manner when they switched from one language to the other. Hence, female and male modulations were not significantly different when speaking in the English sequences: 3.05 semitones for female speakers and 3.08 semitones for male speakers. However, when speaking the sequences in French, female speakers modulated more than male speakers: – female *SD* of 2.94 semitones, male *SD* of 2.46.

Table 4. Average Values of F0 Range in Hertz (Hz) and Semitones (st), Standard Deviation (SD) of F0 in Hertz and Semitones, for Both Female (F) and Male (M) Speakers on Semi-Spontaneous Speech as a Function of the Spoken Language (English or French).

	Semi-spontaneous speech - EN				Semi-spontaneous speech - FR				
Speaker	F0 ran. (Hz)	F0 ran. (st)	SD (Hz)	SD (st)	F0 ran. (Hz)	F0 ran. (dt)	SD (Hz)	SD (st)	% diff. FR/EN SD (dt)
F1	300.32	25.41	42.62	3.94	297.94	25.30	41.65	3.54	**−10.15**
F2	309.61	25.82	34.26	3.24	305.69	25.64	36.92	3.34	**+3.09**
F3	211.16	20.89	22.51	2.23	253.20	23.05	23.44	2.18	**−2.24**
F4	289.14	23.41	28.16	2.52	276.55	23.57	28.38	2.46	**−2.38**
F5	300.88	25.42	40.94	3.64	306.62	25.61	42.90	3.52	**−3.30**
F6	244.63	22.63	30.98	2.75	286.52	24.78	29.70	2.60	**−5.45**
F aver.	**275.96**	**23.93**	**33.25**	**3.05**	**287.75**	**24.66**	**33.83**	**2.94**	**−3.41**
M1	224.44	28.21	34.33	4.21	177.13	28.42	26.78	3.56	**−15.44**
M2	67.52	16.27	7.13	1.58	53.24	12.52	6.87	1.54	**−2.53**
M3	124.70	24.24	21.91	4.36	151.39	21.68	13.37	2.09	**−52.06**
M4	155.71	24.44	20.11	3.18	151.03	24.08	7.53	2.95	**−7.23**
M5	179.88	21.87	25.44	3.21	176.18	21.78	20.98	2.67	**−16.82**
M6	127.53	19.71	11.87	1.94	107.71	17.79	11.75	1.92	**−1.03**
M aver.	**146.63**	**22.46**	**20.13**	**3.08**	**136.11**	**21.05**	**14.55**	**2.46**	**−15.85**

The variation of SD in semitones (expressed in %) between English and French languages for each individual speaker is indicated in the right column.

Vowel Formants

Average formant values in French vowels (F1, F2, and F3) for female and male speakers are shown in Table 5, below. We measured formant values on initial vowels of the dissyllabic words corpus (22 measurements per formant and per vowel for each speaker).

Unsurprisingly, overall formant values appeared to be greater in female speakers. However, the scope of this cross-gender difference strongly varied, depending on the vowel and the formant considered. The female/male ratios were highest on the first formant of the vowel [a] (+39% compared to male values) and on the F2 of the vowel [i] (+31%). On the other hand, on the second formant of the vowel [u], this ratio was very similar values across genders (+5% in female speakers).

We performed a two-factor ANOVA ("*vowel*" and "*speaker's gender*") on F1 values in French. There was a strong and significant main effect of the speaker's gender: $F(1,786) = 459.064$, $p < .0001$. The analysis also detected an interaction between the two factors, showing that the *gender* factor varied with the *vowel*, $F(2,786) = 208.663$, $p < .0001$. For each vowel taken separately, the gender effect factor appeared to be significant, $F(1,262) = 33{,}901$, $p < .0001$ for [u]; $F(1,262) = 351{,}020$, $p < .0001$ for [a]; and $F(1,262) = 79{,}387$, $p < .0001$ for [i].

Similar tests were performed on F2 values. There was a strong and significant overall effect of the speaker's gender, $F(1,786) = 794.071$, $p < .0001$, as well as a significant interaction between the *speaker's gender* and the *vowel* factors, $F(2,786) = 240.866$, $p < .0001$. A closer look at each individual vowel indicated a significant cross-gender difference for [a], $F(1,262) = 319.057$, $p < .0001$ and [i], $F(1,262) = 2050.542$, $p < .0001$, but no significant effect of the *gender* factor on the F2 for the vowel [u], $F(1,262) = 3.333$, $p = .069$.

We performed the same analysis on F3 values and found a very significant main effect of the *speaker's gender,* $F(1,786)=1072.078$, $p < .0001$, and a significant interaction between the *speaker's gender* and the *vowel* factors, $F(2,786) = 38.164$, $p < .0001$. The cross-gender difference was significant for each vowel taken individually: $F(1,262) = 217.589$, $p < .0001$ for [u]; $F(1,262)=357.220$, $p < .0001$ for [a] and $F(1,262) = 604.861$, $p < .0001$ for [i].

Vowel formant values for the English corpus are presented in Table 6 below. Similarly to what was found for the French data, overall formant values were greater among female speakers, but the scope of this female/male difference depended on the vowel and the formant considered. Female/male ratios were highest on the F1 of the vowel [æ] (+37% compared to male values) and on the second formant of the vowel [i:] (+30%). But for this analysis, there was a large cross-gender difference on the second formant of the vowel [u:] (+26% in female speakers), which showed an important contrast with the data obtained on the French vowel counterpart [u].

Table 5. Mean Vowel Formant Frequencies (Hz) for F1, F2 and F3 in French Vowels [i] [a] and [u] Produced by Female (F) and Male (M) Speakers in Read Dissyllabic Words.

		Mean formant frequencies in French vowels (Hz)														
	Speaker	F1	F2	F3	F4	F5	F6	F Avg.	M1	M2	M3	M4	M5	M6	M Avg.	Ratio F/M
[a]	F1	1027	721	696	733	768	791	**789**	602	571	592	563	551	529	**568**	**1.39**
	σ	*47*	*52*	*77*	*36*	*62*	*53*	***124***	*48*	*51*	*72*	*40*	*43*	*36*	***55***	
	F2	1885	1913	1797	1713	1814	1704	**1804**	1547	1584	1606	1569	1453	1581	**1557**	**1.16**
	σ	*134*	*109*	*78*	*91*	*127*	*89*	***131***	*62*	*84*	*71*	*89*	*56*	*97*	***91***	
	F3	2847	2684	2690	2865	2879	2781	**2791**	2248	2479	2521	2445	2563	2615	**2478**	**1.13**
	σ	*66*	*128*	*135*	*53*	*92*	*72*	***124***	*46*	*82*	*124*	*100*	*54*	*84*	***144***	
[u]	F1	278	354	299	326	261	272	**298**	298	248	318	256	267	260	**274**	**1.09**
	σ	*8*	*37*	*16*	*15*	*9*	*16*	***38***	*18*	*6*	*20*	*12*	*14*	*7*	***29***	
	F2	1085	997	1268	920	1250	862	**1064**	959	1021	1112	943	1074	987	**1016**	**1.05**
	σ	*139*	*136*	*361*	*49*	*231*	*71*	***246***	*135*	*172*	*159*	*117*	*218*	*158*	***171***	
	F3	2443	2343	2512	2638	2753	2821	**2585**	2109	2266	2137	1991	2280	2337	**2187**	**1.18**
	σ	*184*	*136*	*135*	*138*	*129*	*94*	***217***	*422*	*55*	*98*	*102*	*111*	*70*	***222***	
[i]	F1	284	349	292	332	262	264	**297**	280	244	300	242	242	250	**260**	**1.14**
	σ	*18*	*48*	*12*	*15*	*13*	*14*	***40***	*11*	*5*	*19*	*8*	*24*	*13*	***27***	
	F2	2783	2504	2572	2677	2619	2752	**2651**	2092	2117	1962	1960	2020	2021	**2029**	**1.31**
	σ	*45*	*127*	*110*	*109*	*46*	*65*	***132***	*53*	*75*	*49*	*57*	*77*	*73*	***87***	
	F3	3815	3733	3395	3593	3446	3537	**3587**	3176	3037	2940	3023	2836	2971	**2997**	**1.20**
	σ	*140*	*183*	*193*	*186*	*140*	*194*	***227***	*57*	*137*	*140*	*131*	*120*	*108*	***156***	

SD are mentioned in italics.

Table 6. Mean Vowel Formant Frequencies (Hz) for F1, F2 and F3 in English Vowels [i:] [æ] and [u:] Produced by Female (F) and Male (M) Speakers in Read Dissyllabic Words.

		Mean formant frequencies in English vowels (Hz)														
	Speaker	F1	F2	F3	F4	F5	F6	F Avg.	M1	M2	M3	M4	M5	M6	M Avg.	Ratio F/M
[æ]	F1	1076	774	721	881	851	861	**861**	627	635	608	724	574	606	**629**	**1.37**
	σ	*51*	*61*	*75*	*66*	*43*	*78*	***128***	*23*	*41*	*47*	*38*	*70*	*33*	***64***	
	F2	1899	1924	1876	1690	1872	1725	**1831**	1561	1575	1538	1546	1554	1617	**1565**	**1.17**
	σ	*134*	*88*	*73*	*66*	*74*	*82*	***126***	*53*	*62*	*35*	*44*	*80*	*50*	***60***	
	F3	2841	2727	2686	2651	2944	2807	**2776**	2302	2388	2530	2471	2515	2547	**2459**	**1.13**
	σ	*80*	*96*	*146*	*121*	*97*	*85*	***145***	*58*	*68*	*67*	*93*	*57*	*69*	***111***	
[u:]	F1	313	334	314	343	291	307	**317**	302	257	347	276	279	293	**292**	**1.08**
	σ	*46*	*21*	*18*	*8*	*8*	*20*	***29***	*15*	*12*	*21*	*13*	*17*	*16*	***32***	
	F2	1218	1598	1495	1638	1601	1281	**1472**	995	1171	1391	987	1249	1234	**1171**	**1.26**
	σ	*187*	*252*	*330*	*279*	*265*	*230*	***304***	*191*	*215*	*208*	*134*	*227*	*228*	***246***	
	F3	2645	2576	2622	2737	2810	2779	**2695**	2280	2227	2198	2104	2259	2339	**2235**	**1.21**
	σ	*80*	*77*	*116*	*114*	*95*	*116*	***131***	*59*	*53*	*79*	*95*	*92*	*73*	***105***	
[i:]	F1	309	317	311	324	290	291	**307**	274	252	341	256	255	276	**276**	**1.11**
	σ	*23*	*35*	*21*	*7*	*10*	*15*	***24***	*11*	*23*	*22*	*26*	*16*	*10*	***36***	
	F2	2792	2675	2592	2873	2633	2804	**2728**	2165	2147	2029	2110	2048	2100	**2100**	**1.30**
	σ	*95*	*138*	*107*	*75*	*80*	*101*	***142***	*50*	*57*	*93*	*61*	*97*	*63*	***86***	
	F3	3704	3818	3256	3394	3245	3386	**3467**	3074	2998	2760	3028	2703	2757	**2887**	**1.20**
	σ	*164*	*225*	*103*	*116*	*96*	*251*	***275***	*84*	*146*	*123*	*135*	*111*	*95*	***189***	

SD are mentioned in italics.

We conducted a two-factor ANOVA ("*speaker's gender*" and "*vowel*") on F1 values in English. Overall, it showed a strong and significant effect of the speaker's gender, $F(1,786) = 451.643$, $p < .0001$. This test also revealed an interaction between the two factors *gender* and *vowel*, $F(2,786) = 227.075$, $p < .0001$. For each individual vowel, the effect of the *gender* factor appeared to be significant: $F(1,262) = 42.205$, $p < .0001$ for [u:]; $F(1,262)=347.613$, $p < .0001$ for [æ]; and $F(1,262) = 68.363$, $p < .0001$ for [i:].

We conducted similar statistical tests for F2 values, and these results revealed a very strong and significant overall effect of the *speaker's gender*, $F(1,786) = 942.947$, $p < .0001$, as well as a significant interaction between the two factors, $F(2,786) = 240.866$, $p < .0001$. Contrary to the French data, the cross-gender difference appeared to be significant for each vowel taken separately: for the back vowel [u], $F(1,262) = 78.112$, $p < .0001$; for [æ], $F(1,262) = 481.944$, $p < .0001$; and for [i:], $F(1,262) = 1886.871$, $p < .0001$.

The same analysis was performed for F3, and results showed a significant global effect of the *speaker's gender*, $F(1,786) = 1406.561$, $p < .0001$, and a significant interaction between the *speaker's gender* and the *vowel* factors, $F(2,786) = 39.889$, $p < .0001$. A closer examination of each individual vowel indicated a significant cross-gender difference for [u], $F(1,262) = 990.112$, $p < .0001$, for [æ], $F(1,262) = 398.715$, $p < .0001$, and for [i], $F(1,262) = 399.050$, $p < .0001$. In order to make these results clearly visible, we present F1 and F2 values in English and French vowels in a vowel chart (see Figure 2), using a template from SaRP software (Nikolov et al., 2011).

VOT

Average voice onset times (ms) in the French language are presented in Table 7. They were measured on initial stop consonants of the dissyllabic words corpus (6 occurrences per consonant for each speaker). In voiceless stop consonants, VOT values appeared to be greater among female speakers. This was true for

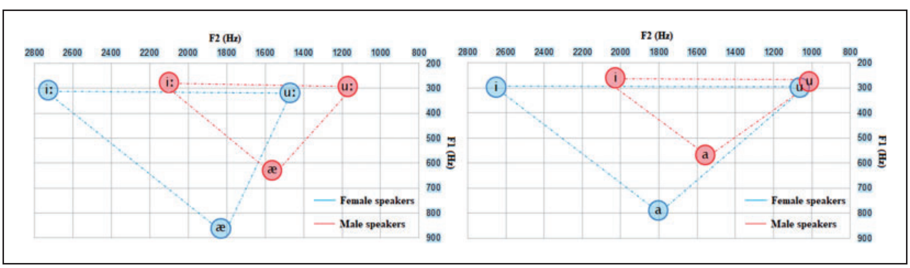

Figure 2. Vowel Chart Representing Mean Vowel Formant Frequencies (Hz) for English Vowels [i:] [æ] and [u:] (on the Left), and French Vowels [i], [a], [u] (on the Right) Produced by Female (in Blue) and Male (in Red) Bilingual Speakers.

Table 7. Average VOT Values (ms) for Female (F) and Male (M) Speakers in French Dissyllabic Words (6 Occurrences per Consonant per Speaker).

	VOT (ms) in French							
Speaker	[p]	[t]	[k]	Voiceless avg.	[b]	[d]	[g]	Voiced avg.
F1	58	64	89	**70**	−89	−69	−91	**−83**
F2	16	32	41	**30**	−73	−72	−72	**−72**
F3	53	64	64	**60**	−53	−34	−89	**−58**
F4	44	58	61	**55**	−26	−70	−40	**−45**
F5	44	63	70	**59**	−79	−83	−83	**−82**
F6	17	40	56	**38**	−84	−74	−84	**−81**
F average	**39**	**53**	**64**	**52**	**−67**	**−67**	**−76**	**−70**
M1	24	38	42	**34**	−115	−91	−91	**−99**
M2	34	57	57	**50**	−85	−59	−64	**−70**
M3	36	48	59	**47**	−46	−38	−34	**−39**
M4	15	32	43	**30**	−80	−49	−35	**−55**
M5	44	56	58	**53**	−52	−51	−58	**−53**
M6	36	46	53	**45**	−71	−86	−26	**−61**
M average	**31**	**46**	**52**	**43**	**−75**	**−62**	**−51**	**−63**

each consonant ([p], [t] and [k]). Considering all voiceless consonants collectively, VOTs were 21% higher among female speakers (+9ms). Regardless of the speaker's gender, these values followed universal tendencies for VOT: [p] < [t] < [k].

Voiced consonants all presented negative values due to pre-voicing, a typical characteristic of French voiced plosives (Pépiot, 2016). Overall, this pre-voicing was longer among female than among male speakers, with seven additional milliseconds for females (+11%).

We conducted a two-factor ANOVA ("*gender*" and "*consonant*") on VOT for French voiceless plosives. This analysis showed a significant effect of the *gender* factor, $F(1,210) = 12.120$, $p < .001$, as well as the *consonant* factor, $F(2,210) = 27.981$, $p < .0001$. However, there was no interaction between the two factors, confirming that females' VOT values tended to be significantly longer in voiceless plosives, regardless of the consonant ([p], [t] or [k]).

We also performed a similar statistical analysis on French voiced plosives. The two-factor ANOVA then showed that, despite an apparent cross-gender difference, the influence of the *gender* factor did not reach statistical significance, $F(1,210) = 1.869$, $p = .1730$. This may be explained by a relatively small number of occurrences. The *consonant* factor was also statistically insignificant, $F(2,210) = 0.683$, $p = .5064$, and no significant interaction was found between the two factors.

Table 8. Average VOT Values (ms) for Female (F) and Male (M) Speakers in English Dissyllabic Words (6 Occurrences per Consonant for Each Speaker).

	VOT (ms) in English							
Speaker	[p^h]	[t^h]	[k^h]	Voiceless avg.	[b]	[d]	[g]	Voiced avg.
F1	81	93	116	**96**	11	14	29	**18**
F2	32	58	62	**51**	−83	−92	−88	**−88**
F3	65	76	81	**74**	10	16	25	**17**
F4	46	65	77	**63**	11	20	24	**18**
F5	65	80	87	**77**	7	20	34	**20**
F6	62	80	100	**81**	7	17	23	**15**
F average	**58**	**75**	**87**	**74**	**−6**	**−1**	**8**	**0**
M1	42	61	57	**53**	−46	−14	−59	**−40**
M2	40	56	64	**53**	6	19	24	**16**
M3	47	64	71	**61**	11	21	28	**20**
M4	45	50	71	**55**	11	15	26	**17**
M5	45	65	56	**55**	−16	−35	15	**−12**
M6	60	67	78	**68**	−7	11	9	**4**
M average	**46**	**61**	**66**	**58**	**−7**	**3**	**7**	**1**

Average voice onset times (ms) measured on English dissyllabic words are shown in Table 8, below (six occurrences per consonant per speaker). Similarly to what was found for participants speaking in French, VOT values for voiceless plosives appeared to be higher among female speakers, regardless of the consonant. This tendency was even accentuated when speaking in English, since VOTs in voiceless consonants were, on average, 28% higher among female speakers (+16ms). These values also followed universal VOT tendencies for both genders.

In voiced plosives, female and male speakers exhibited fairly similar values. The average figure, when all consonants were taken together, was around 0 milliseconds for both genders. These data are consistent with typical findings among monolingual speakers (Pépiot, 2016; Whiteside & Irving, 1997), except that two of these participants, *F2* and *M1*, systematically produced some prevoicing, clearly indicating an interference of French as an L2.

We performed a two-factor ANOVA ("*gender*" and "*consonant*") on VOT values for English voiceless plosives. This analysis showed a strong and significant effect of *gender,* $F(1{,}210) = 49.491$, $p < .0001$, as well as of *consonant,* $F(2{,}210) = 39.777$, $p < .0001$. We found no significant interaction between the two factors. This confirms that, similarly to French, females' VOT values in English tended to be significantly greater than males' in voiceless plosives, regardless of the consonant. We conducted the same statistical test on English voiced plosives and found no significant effect of the *speaker's gender,* $F(1{,}210) = .030$, $p = .8628$

Table 9. Average VOT Contrast (ms) Between Voiced and Voiceless Plosives for Female (F) and Male (M) Speakers in Dissyllabic Words, as a Function of the Spoken Language (French or English).

	VOT contrast in French (ms)				VOT contrast in English (ms)			
Speaker	[p] vs. [b] mean	[t] vs. [d] mean	[k] vs. [g] mean	Voiced vs. voiceless avg.	[p^h] vs. [b] mean	[t^h] vs. [d] mean	[k^h] vs. [g] mean	Voiced vs voiceless avg.
F1	147	132	180	**153**	70	79	87	**78**
F2	89	104	113	**102**	116	150	150	**138**
F3	106	98	152	**119**	55	60	57	**57**
F4	70	128	101	**100**	35	45	53	**44**
F5	124	146	154	**141**	58	61	54	**57**
F6	101	115	140	**119**	56	64	78	**66**
F avg.	**106**	**120**	**140**	**122**	**65**	**76**	**80**	**74**
M1	138	128	133	**133**	87	75	117	**93**
M2	120	116	121	**119**	34	37	41	**37**
M3	81	86	93	**87**	37	44	43	**41**
M4	95	82	78	**85**	35	36	45	**38**
M5	95	107	116	**106**	60	99	41	**67**
M6	107	132	79	**106**	67	57	69	**64**
M avg.	**106**	**109**	**103**	**106**	**53**	**58**	**59**	**57**

on these data. Moreover, we found no significant effect of the *consonant,* $F(2{,}210) = 2.415$, $p = .0919$, and no interaction between these factors.

The VOT contrast (in ms), corresponding to the difference between the VOT of a voiceless plosive and the VOT of its voiced counterpart, is presented in Table 9 for both languages. Overall, the VOT contrast appeared to be greater in female speech. This phenomenon was found in both languages but was more pronounced in English, where female speakers exhibited an average VOT contrast 30% longer than did males (15% in French).

In order to test if those cross-gender differences were statistically significant, we performed a two-factor ANOVA ("*gender*" and "*consonant type*") on the data. In French, there was a significant effect of the *speaker's gender,* $F(1{,}102) = 5.173$, $p < .01$, but no effect of the *consonant type,* $F(2{,}102) = 1.568$, $p = .2134$, and there was no interaction between the two factors.

Similar results were found on English consonants, with a significant effect of the *gender* factor, $F(1{,}102) = 7.159$, $p < .01$, no significant effect of the *consonant type,* $F(2{,}102) = 1.035$, $p = .3588$, and no significant interaction. This result shows that in both languages, VOT contrast tended to be significantly longer in female than in male speech, regardless of the consonant type. This VOT contrast reinforcement might be an indication of the participants' search for intelligibility.

Table 10. Average H1-H2 Values in Decibels (dB) for Female (F) and Male (M) Speakers in Open Vowels (11 X 2 Occurrences) as a Function of the Spoken Language (English or French).

	H1-H2 mean in open vowels (dB)		
Speaker	English	French	FR-EN diff.
F1	7,60	7,77	**0,16**
F2	2,75	1,58	**−1,17**
F3	7,48	5,62	**−1,86**
F4	4,50	4,30	**−0,20**
F5	10,13	9,66	**−0,47**
F6	4,71	4,66	**−0,05**
F average	**6,20**	**5,60**	**−0,60**
M1	−3,96	−2,52	**+1,45**
M2	−1,59	−0,08	**+1,51**
M3	1,52	2,42	**+0,90**
M4	1,54	2,07	**+0,54**
M5	6,85	9,09	**+2,25**
M6	0,53	0,75	**+0,22**
M average	**0,81**	**1,96**	**+1,14**

The absolute difference between English and French languages for each individual speaker is indicated in the right column.

Phonation Type

Mean intensity differences (dB) between the first harmonic (H1) and the second harmonic (H2) in both languages and as a function of speaker's gender is presented in Table 10, below. We measured this difference on open vowels, yielding a total of 22 measurements per speaker in each language (11 words contained vowel [a] in the French corpus while 11 words contained vowel [æ] in the English corpus, and they were all pronounced twice). The relative strength of H1 was correlated with glottal open quotient (GOQ): the stronger H1 was, the higher the GOQ. Thus, a high H1-H2 value indicated a *breathy voice*, while negative values were associated with a *creaky voice*. H1-H2 intensity differences appeared to be greater in female speakers in both languages. This cross-gender difference reached 5.39 dB in English and 3.64 dB in French. Overall, female speakers exhibited smaller H1-H2 values in French, whereas male speakers presented higher values in French.

We performed a one-factor ANOVA ("*gender*") on the data obtained in the French language. This analysis revealed a significant effect of gender on H1-H2 intensity differences, $F(1,262) = 129.133$, $p < .0001$. We conducted the same statistical test on the English data. Similarly to what was found in French, this test indicated a significant effect of speaker's gender, $F(1,262) = 61.273$, $p < .0001$. These results confirmed that in both languages the H1-H2 difference was significantly higher among female than male speakers, suggesting that female speakers used a more breathy phonation type than did male speakers.

In order to test whether cross-language variations were significant, we performed a two-factor ANOVA ("*spoken language*" and "*gender*") on the whole set of data (French and English). The analysis showed a significant interaction between the two factors, $F(1,524) = 6.871$, $p < .01$, confirming that the adaptation of phonation type when switching from one language to the other was gender-dependent, hence lowering H1-H2 values in French for female speakers and increasing them for male speakers so as to result in a smaller cross-gender difference in this language.

Discussion

This multiparametric study led to many interesting results that need to be discussed and put into perspective. Regarding the speakers' use of the fundamental frequency, we found a significant interaction between the factors "*spoken language*" and "*gender*" on average F0 in read speech and on F0 modulation in spontaneous speech such that more females than males produced higher average F0 in French than English. We observed a similar but non-significant trend for average F0 in spontaneous speech as well. This interaction, indicating that, in these contexts, both language and gender determine average F0 and F0 modulations in complex patterns, can be interpreted as a greater cross-linguistic gender difference in average F0 use when speaking in French versus English.

For speech modulations, F0 *SDs,* expressed in semitones, were generally reduced for all participants when they spoke in French, relative to English; and this decreased modulation was accentuated among males – 15.61% in read speech and 15.85% in spontaneous speech versus females – 14.36% in read speech and only 3.41% in spontaneous speech. Moreover, we noticed that, during the semi-spontaneous sequences produced in English, male speakers modulated as much as female speakers, consistent with Henton's (1995) finding of no significant differences between female and male American speakers when F0 modulations where expressed in semitones. Our finding that male speakers modulated less than female speakers when they spoke in French versus English confirmed Pépiot's (2014b) suggestion that F0 modulations in French were gender-dependent. Later, in 2017, Pépiot found that, for French (but not American listeners), F0 modulation was a relevant acoustic cue for identifying a speaker's gender (more modulation was associated with female voices).

Regarding vowel formants (F1, F2 and F3) in dissyllabic read words, we found, as expected, that formant frequencies were globally higher in female speakers in both languages. The biggest cross-gender differences in English were found on the F1 of the vowel [æ] (+37%) and on the F2 of [i:] (+30%). In French, these differences were maximal on the F1 of [a] (+39%) and on the

F2 of [i] (+31%). This could suggest that cross-gender differences on this acoustic parameter are fairly similar across languages.

We observed a striking phenomenon on the back vowels [u] (in French) and [u:] (in English). While in English, a very large cross-gender difference was evident on the F2 (+26% in female speakers), there was no significant difference in French (+5% only in female speakers). This finding is consistent with Pépiot (2015). The traditional morphological explanation for cross-gender differences in vowel formant values (based on vocal tract length) was clearly irrelevant. Rather, gender-dependant vocal practices may be relevant. Female speakers might have produced a greater protrusion of the lips when pronouncing the French vowel [u] in order to reach a low F2 frequency (Fant, 1966, 1975). In contrast, they may have produced a quite central position of the tongue when pronouncing the English vowel [u:], hence obtaining high F2 values.

Regarding the production of stop consonants, our results showed that VOT was significantly longer for female speakers on voiceless stops in both languages, as their values were 21% higher than males' in French, and 28% higher in English. These results confirm previous findings among English monolingual speakers (Robb et al., 2005; Swartz, 1992; Whiteside & Irving, 1997) and French monolingual speakers (Pépiot, 2016). Regardless of the speaker's gender, these values followed universal VOT tendencies ([p] < [t] < [k]). On voiced plosives, the close to zero values were very similar for both genders in English; there were no statistically significant gender differences in French either, though there was a tendency for female French speakers to exhibit longer pre-voicing than males (+11% on average).

VOT *contrast*, which corresponds to the difference between the VOT of a voiceless stop and the VOT of its voiced equivalent, was significantly longer among female speakers in both English and French, with a greater cross-gender difference in English (+30%) than French (+15%). This phenomenon can be seen as an illustration of a broader gender-related socio-phonetic tendency in that using a greater VOT contrast might be a way for female speakers to increase *speech intelligibility*, which is recurrent trend on several acoustic parameters (see Pépiot, 2015, 2016; Simpson, 2009).

H1-H2 measurements in open vowels gave valuable indications of speakers' phonation type. In both French and English, significant cross-gender differences were found in that female voices exhibited higher H1-H2 values than males. This suggests that females tended to speak with a breathier voice quality. However, a significant interaction between "spoken language" and "speaker's gender" showed that the adaptation of phonation type when switching from one language to the other was gender-dependent. Female speakers tended to use slightly higher H1-H2 values in English, while male speakers presented lower values than in French. This resulted in a wider cross-gender difference in English. These results support the claim that female/male differences in phonation type are a socio-phonetic and language-dependant phenomenon (Henton, 1989a;

Pépiot, 2014b). They are also consistent with what was found in perceptual experiments such as a cross-language study (Pépiot, 2017) showing that phonation type was an important acoustic cue for American (but not for French) listeners for identifying a speaker's gender from their speech.

Limitations and Directions for Future Research

A main limitation of this study was its small number of participants. In order to generalize these findings to other populations it will be necessary to replicate this study with a larger number and diversity of participant speakers. Furthermore, we analyzed speech produced by bilingual native English speakers whose second language was French. Even though our results are very consistent with previous findings among French and English monolinguals, it would be interesting to determine whether these results can be replicated with French speakers whose second language is English. The increase in F0 among females speaking French could be partly due to the stress induced by speaking a second language, since stress has been found to induce an F0 increase (Scherer, 1986). Eventually, this experiment might also be extended to female and male speakers of other languages, in order to study gender-dependant vocal practices in different cultural regions.

Conclusion

Overall, the analysis of these acoustic parameters within a cross-sectional study of *language* and *gender* in the speech of a small sample of bilingual participants has brought to light new facts not previously explored in prior research (e.g., Altenberg & Ferrand, 2006; Lee & Van Lancker Sidtis, 2017; Mennen et al., 2012). Specifically, the current study showed that *the process by which female or male voices are produced involves different vocal practices, and these practices vary from one language to another*. Therefore, these data confirm that vocal acoustic parameters, such as F0 and vowel formants, are not *essential* characteristics of speakers that depend only on their anatomy. These acoustic parameters also depend on the way speakers *use* their vocal apparatus. During the socialization process, in which gender norms and ideologies are internalized and incorporated, speakers learn to use their body, hence also their vocal apparatus, to accomplish social practices through which they are identified as members of a specific gender group (Bourdieu, 1972, 1998). They learn to move, dress, and to speak in ways that performatively constitutes them as women or men (Butler, 1990). For example, a study conducted by Perry et al. (2001) showed that girls and boys start at a very early age – around 4 – to speak differently. Since prepubescent children don't present significative gender differences in the morphology of their vocal apparatus, this difference in speech is the sole result of a gender distinctive way to use the vocal apparatus. In the same manner, adult speakers tend to exaggerate female-male differences that are due to biological differences (Sachs, 1975, p. 154).

This constitutes a further argument to abandon simplistic understandings of the relationship between voice and anatomy and adopt, instead, a model in which social factors, such as gender dynamics, are given more consideration in phonetic studies.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) received no financial support for the research, authorship, and/or publication of this article.

ORCID iD

Erwan Pépiot https://orcid.org/0000-0001-5457-1384

References

Abitbol, J., Abitbol, P., & Abitbol, B. (1999). Sex hormones and the female voice. *Journal of Voice: Official Journal of the Voice Foundation*, *13*(3), 424–446.

Altenberg, E. P., & Ferrand, C. T. (2006). Fundamental frequency in monolingual English, bilingual English/Russian, and bilingual English/Cantonese young adult women. *Journal of Voice: Official Journal of the Voice Foundation*, *20*(1), 89–96.

Arnold, A. (2015). Voix et transidentité: Changer de voix pour changer de genre [Voice and transidentity: Changing your voice to change your gender]. *Langage et Société*, *151*(1), 87–105.

Arnold, A. (2016). Voix. In J. Rennes (Ed.), *Encyclopédie critique du genre [Critical encyclopedia of gender]* (pp. 713–721). La Découverte.

Austin, W. M. (1965). Some social aspects of paralanguage. *Canadian Journal of Linguistics/Revue Canadienne de Linguistique*, *11*(1), 31–39.

Boë, L.-J., Contini, M., & Rakotofiringa, H. (1975). Étude statistique de la fréquence laryngienne [Statistical study of laryngeal frequency]. *Phonetica*, *32*(1), 1–23.

Boersma, P., & Weenink, D. (2017). Praat: Doing phonetics by computer [Computer Program]. Version 6.0.36. www.praat.org

Bourdieu, P. (1972). *Esquisse d'une théorie de la pratique* [Outline of a theory of practice]. Droz.

Bourdieu, P. (1998). *La domination masculine*. Seuil.

Butler, J. (1990). *Gender trouble: Feminism and the subversion of identity*. Routledge.

Fant, G. (1966). A note on vocal tract size factors and non-uniform F-pattern scaling. *Speech Transmission Laboratory, Quarterly Progress and Status Report*, *81*, 21–38.

Fant, G. (1975). Non-uniform vowel normalization. Speech transmission laboratory. *Quarterly Progress and Status Report*, *16*, 1–19.

Gordon, M., & Ladefoged, P. (2001). Phonation types: A crosslinguistic overview. *Journal of Phonetics*, *29*(4), 383–406.
Grosjean, F., & Li, P. (2013). *The psycholinguistics of bilingualism*. Wiley-Blackwell.
Henton, C. (1989a). Fact and fiction in the description of female and male pitch. *Language & Communication*, *9*(4), 299–311.
Henton, C. (1989b). Sociophonetic aspects of creaky voice. *Journal of the Acoustical Society of America*, *86*(S1), S26–S26.
Henton, C. (1995). Pitch dynamism in female and male speech. *Language & Communication*, *15*, 43–61.
Henton, C., & Bladon, R. (1985). Breathiness in normal female speech: Inefficiency versus desirability. *Language & Communication*, *5*(3), 221–227.
Hirst, D., Di Cristo, A., & Nishinuma, Y. (2001). Prosodic parameters of French: A cross-language approach. In *Contrastive studies of Japanese and other languages series* (pp. 7–20). National Institute for Japanese Language.
Honjo, I., & Isshiki, N. (1980). Laryngoscopic and voice characteristics of aged persons. *Archives of Otolaryngology*, *106*, 14–150.
Johnson, K. (2005). Speaker normalization in speech perception. In D. Pisoni & R. Remez (Eds.), *The handbook of speech perception* (pp. 363–389). Blackwell Publishers.
Kahane, J. C. (1978). A morphological study of the human prepubertal and pubertal larynx. *The American Journal of Anatomy*, *151*(1), 11–19.
Klatt, D. H., & Klatt, L. C. (1990). Analysis, synthesis and perception of voice quality variations among female and male talkers. *The Journal of the Acoustical Society of America*, *87*(2), 820–857.
Karlsson, F., Zetterholm, E. & Sullivan, K. (2004). Development of a gender difference in voice onset time. In *Proceedings of the 10th Australian International Conference on Speech Science & Technology*, Sydney, 316–321.
Lakoff, R. (1975). *Language and woman's place*. Harper & Row.
Lee, B., & Van Lancker Sidtis, D. (2017). The bilingual voice: Vocal characteristics when speaking two languages across speech tasks. *Speech, Language and Hearing*, *20*(3), 174–185.
Matar, N., Portes, C., Lancia, L., Legou, T., & Baider, F. (2016). Voice quality and gender stereotypes: A study of Lebanese women with Reinke's edema. *Journal of Speech, Language, and Hearing Research*, *59*(6), S1608–S1617.
Mennen, I., Schaeffler, F., & Docherty, G. (2012). Cross-language differences in fundamental frequency range: A comparison of English and German. *The Journal of the Acoustical Society of America*, *131*(3), 2249–2260.
Nikolov, R., Ditcheva, M., Dommergues, J.-Y., Herry-Bénit, N. & Ryst, E. (2011). *SaRP: Speech Analyzer Rapid Plot (2006-2011) Plotting vowels in F2-F1 scatter charts with multiple data sets*. http://web.uni-plovdiv.bg/rousni/sarp
Oh, E. (2011). Effects of speaker gender on voice onset time in Korean stops. *Journal of Phonetics, 39*, 59–67.
Pépiot, E. (2014a). Voix et genre: Un état de la question [Voice and gender: A state of the art]. In A. H. Ibrahim (Eds.), *La langue, la voix, la parole* (pp. 53–86). CRL.
Pépiot, E. (2014b). *Male and female speech: A study of mean F0, F0 range, phonation type and speech rate in Parisian French and American English speakers* [Paper presentation]. 7th International Conference on Speech Prosody (pp. 305–309). https://www.research

gate.net/publication/315880083_Proceedings_of_the_7th_International_Conference_on_Speech_Prosody

Pépiot, E. (2015). Voice, speech and gender: Male-female acoustic differences and cross-language variation in English and French speakers. Corela, HS-16. https://doi.org/10.4000/corela.3783

Pépiot, E. (2016). Voix de femmes, voix d'hommes: Une étude du voice onset time, de la répartition consonnes/voyelles et du débit de parole chez des locuteurs francophones et anglophones américains [Female voices, male voices: A study of voice onset time, consonant/vowel distribution and speech rate among French-speaking and English-speaking American speakers]. *Actes de la Conférence JEP-TALN-RECITAL, 2016*, 759–767.

Pépiot, E. (2017). *Gender identification from speech in Parisien French and American English speakers* (pp. 60–72). Paisii Hilendarski University of Plodiv, Bulgaria – Research Papers, 55.

Perry, T. L., Ohde, R. N., & Ashmead, D. H. (2001). The acoustic bases for gender identification from children's voices. *The Journal of the Acoustical Society of America*, *109*(6), 2988–2998.

Robb, M., Gilbert, H., & Lerman, J. (2005). Influence of gender and environmental setting on voice onset time. *Folia Phoniatrica et Logopaedica : Official Organ of the International Association of Logopedics and Phoniatrics (Ialp)*, *57*(3), 125–133.

Rose, P. (1991). How effective are long term mean and standard deviation as normalization parameters for tonal fundamental frequency? *Speech Communication*, *10*(3), 229–247.

Sachs, J. (1975). Cues to the identification of sex in children's speech. In B. Thorne & N. Henley (Éds.), *Language and sex : Difference and dominance*. Newbury House Publishers.

Scherer, K. R. (1986). *Voice, stress, and emotion. Dynamics of stress* (pp. 157–179). Springer.

Simpson, A. P. (2009). Phonetic differences between male and female speech. *Language and Linguistics Compass*, *3*(2), 621–640.

Swartz, B. L. (1992). Gender difference in voice onset time. *Perceptual and Motor Skills*, *75*(3), 983–992.

Traunmüller, H., & Eriksson, A. (1995). *The frequency range of the voice fundamental in the speech of male and female adults*. http://www2.ling.su.se/staff/hartmut/f0_m&f.pdf

Van Bezooijen, R. (1995). Sociocultural aspects of pitch differences between Japanese and Dutch women. *Language and Speech*, *38*(3), 253–265.

Whiteside, S. P., & Irving, C. J. (1997). Speakers' sex differences in voice onset time, some preliminary findings. *Perceptual and Motor Skills*, *85*(2), 459–463.

Wolk, L., Abdelli-Beruh, N. B., & Slavin, D. (2012). Habitual use of vocal fry in young adult female speakers. *Journal of Voice: Official Journal of the Voice Foundation*, *26*(3), e111–e116.

Author Biographies

Erwan Pépiot, PhD, is a Senior Lecturer in phonetics at the University of Paris 8. His reaseach focuses on sociophonetics, especially on the relationship between voice, speech and gender.

Aron Arnold, PhD, is a scientific collaborator of the VALIBEL Research Center at the Catholic University of Louvain. His research focuses on the sociophonetics of gender and sexuality.