ПЛОВДИВСКИ УНИВЕРСИТЕТ „ПАИСИЙ ХИЛЕНДАРСКИ“ – БЪЛГАРИЯ
НАУЧНИ ТРУДОВЕ, ТОМ 55, КН. 1, СБ. Б, 2017 – ФИЛОЛОГИЯ,
PAISII HILENDARSKI UNIVERSITY OF PLOVDIV – BULGARIA
RESEARCH PAPERS, VOL. 55, BOOK 1, PART B, 2017 – LANGUAGES AND LITERATURE

GENDER IDENTIFICATION FROM SPEECH IN PARISIAN FRENCH AND AMERICAN ENGLISH SPEAKERS

Erwan Pépiot
University Paris 8

The present study focuses on gender identification obtained from a speech experiment. It was conducted jointly on 25 Parisian French native listeners with French stimuli, and 25 American English listeners with English stimuli. Extracts from (C)VCV words and pseudo-words were presented to the participants, with the use of the gating paradigm. The listeners had to identify the speaker's gender and indicate their degree of certainty. In both languages, the percentages of correct identification were significantly above chance for initial voiceless consonants, and close to 100 % with initial vowels. An acoustic analysis performed on the stimuli showed that American English and French listeners did not use identical strategies: mean f0 and voice quality (H1-H2) seemed to have more influence on American English listeners' judgements than on those of French listeners who appeared to rely more on vowel formant frequencies and f0 range.

Key words: *Gender identification from speech, cross-gender acoustic differences, cross-language variations, Parisian French listeners, American English listeners.*

Introduction

A large number of perceptual studies were conducted in order to test the ability of listeners to identify a speaker's gender from his/her speech. When using natural (i.e. non synthetized) speech, listeners get between 63 and 90 % correct gender identification from voiceless fricatives, depending on stimulus length (Schwartz, 1968). With isolated vowels, scores vary from 96 % (Lass & al., 1976) to 98.9 % (Whiteside, 1998a). In twelve syllable sentences, percentages of correct identifications almost reach 100 % (Pépiot, 2011). However, these studies were conducted with different methods and at different times. The use of the gating paradigm, in which items are presented in segments of increasing duration (Grosjean, 1980), could provide very interesting data if used in a gender identification from speech experiment.

The relative influence of acoustic parameters in gender identification from speech is still much debated. Vowel formant frequencies were shown to be relevant clues (Arnold, 2012), as well as mean f0 (Coleman, 1976, Pépiot, 2011). Consonant noise frequency also appears to be an important parameter, at least in voiceless fricatives (Schwartz, 1968; Whiteside, 1998b). According to a majority of authors who conducted their experiments on American English or British English listeners, mean f0 is the most salient cue.

However, a recent experimental study led on Parisian French listeners (Arnold, 2012), indicates that resonant frequencies and particularly vowel formant frequencies, are more important than mean f0 in influencing listeners' judgments.

How can one account for such contradictory results? It is known that cross-gender acoustic differences slightly vary from one language to another (Johnson, 2005; Pépiot, 2014b). Listeners' strategies to identify gender from speech could then be language-dependant, thus different in English and French listeners.

It then appears relevant to perform a cross-language gender identification from speech experiment using natural speech and applying the gating paradigm. This study will be conducted jointly on American English listeners, with English stimuli, and Parisian French speakers, with French stimuli. The influence of various acoustic parameters shall be measured through correlation tests. The following research questions will be investigated: (1) What amount of speech is necessary for a listener to correctly identify a speaker's gender? (2) Which acoustic parameters are used by listeners for this task? (3) Are some acoustic segments more favorable than others for gender identification? (4) Are listener's strategies language-dependent?

1. Material and method

1.1 Linguistic material

French and English linguistic material was required for the experiment. We decided to use disyllabic words and pseudo-words. Their selection was based on two main criteria: make the two corpora as similar as possible (e.g. English inter-dental fricatives were dismissed as there is no equivalent in French), and limit the number of combinations by choosing only the most relevant phonemes (e.g. cardinal vowels) while holding constant the last CV sequence: /pi/ was chosen as it can appear on word final position in both languages. Twenty-seven (C)VCV words or pseudo-words were finally

chosen for each language (e.g. /tipi/, /tapi/, /tupi/ for the French corpus; /ˈti:pi/ , /ˈtæpi/, /ˈtu:pi/ for the English corpus).

English words are to be read by American English speakers and used for the experiment conducted on American English listeners, while French words are to be read by French speakers and used for the experiment conducted on French listeners.

1.2 Speakers

Eight monolingual speakers were recorded for the experiment. Four of them were native Parisian French speakers (2 women, 2 men) and four others were native Northeastern American English speakers (2 women and 2 men). Two more speakers in each language (1 woman and 1 man) were recorded exclusively to obtain warm-up items. Speakers were aged from 23 to 40, non-smokers with no reported speech disorder.

1.3 Recording procedure

Recordings took place in a quiet room, using a digital recorder *Edirol R09-HR* by Roland. English speakers read the English corpus aloud and French speakers the French one. Words were presented to the participants in an orthographical transcription. Moreover, in order to make prosodic parameters consistent, words were placed into a frame sentence: "He said "*WORD*" twice" for the English corpus and "Il a dit "*MOT*" deux fois" for the French one. Speakers were asked to say each sentence twice, at a normal speech rate.

1.4 Acoustic Analysis

An acoustic analysis was performed on the stimuli, by using *Praat*. Duration, mean f0 and f0 range (in semitones) of entire words were measured. We also measured duration, mean f0, F1, F2 and F3 frequencies in first syllable's vowels, that will be referred to as V1. H1-H2 values (dB) were taken to get an indication of phonation type, but only in open vowels: such measurement would not be relevant in other vowel types (Klatt & Klatt, 1990). The detailed results of this acoustic analysis are presented in Pépiot (2016a, 2016b).

1.5 Listeners

Two groups of listeners took part in the experiment: a group of American English speakers, and a group of Parisian French speakers:

– **Parisian French listeners group**: 25 participants, 17 women and 8 men, aged from 18 to 48 years old. Mean age: 24.2 (21.9 for women and 29.1 for men).

– **American English listeners group**: 25 participants, 18 women and 7 men, aged from 18 to 38 years old. Mean age: 24.8 (24.3 for women and 26.1 for men).

1.6 Experimental procedure

In a gender identification from speech experiment, audio stimuli are presented to the participant who has to tell if the speaker is a woman or a man. For the current study, we used the gating paradigm (Grosjean, 1980): thus, the length of the stimuli will be increasing as the experiment goes on. For the current study, we used 4 groups of stimuli:

– **96 items containing C1** ([k], [t], [s], etc.).
– **12 items containing V1** ([i], [a], [u], etc.), from VCV words.
– **96 items containing C1 and V1** ([ki], [ta], [tu], etc.).
– **108 items with entire words** ([kipi], [tapi], [tupi], etc.)

Stimuli were played following this order. Inside the four groups of stimuli, the order was randomized, thus different for each participant.

The experiment was performed on a computer, by using software Perceval (André & al., 2003) and took place in a quiet room. For each stimulus, participants had to choose between "female speaker" or "male speaker" and to indicate a degree of certainty on a scale going from 0 (not sure at all) to 7 (completely sure).

2. Results

2.1 All stimuli

Results for French listeners in the five main types of stimuli are presented in Table 2. Percentages of correct gender identifications appear to be relatively high, even in voiceless initial consonants. For each of the five categories, a one-tailed *t*-test reveals that the scores are significantly above chance ($p < 0.0001$), which is here at 50 %.

Table 1 Percentages of correct gender identification and mean degree of certainty for the 5 main types of stimuli in French listeners

Stimulus type	N of items	N of correct answers	N of incorrect answers	Percentage of correct identification	Mean degree of certainty
Voiceless initial C	1200	788	412	**65.67**	**3.56**
Voiced initial C	1200	1134	66	**94.50**	**4.76**
Initial V	300	294	6	**98.00**	**6.54**
C + V	2400	2358	42	**98.25**	**6.64**
CVCV word	2700	2698	2	**99.93**	**6.92**
All types	**7800**	**7272**	**528**	**93.23**	**5.97**

Results for American English listeners in the five main types of stimuli are detailed in Table 2. As for French listeners, percentages of correct gender identifications are quite high, even in voiceless consonants. A one-tailed *t*-test shows that scores are significantly above chance level in each of the five categories of stimuli ($p < 0.0001$ in all cases).

Table 2 Percentages of correct gender identification and mean degree of certainty for the five main types of stimuli in American English listeners

Stimulus type	N of items	N of correct answers	N of incorrect answers	Percentage of correct identification	Mean degree of certainty
Voiceless initial C	1800[a]	1377	423	**76.50**	**2.85**
Voiced initial C	600	596	4	**99.33**	**6.00**
Initial V	300	293	7	**97.67**	**6.33**
C + V	2400	2364	36	**98.50**	**6.39**
CVCV word	2700	2684	16	**99.41**	**6.54**
All types	**7800**	**7314**	**486**	**93.77**	**5.60**

[a] English plosives /d/ and /g/ were phonetically voiceless in this context.

2.2 CV sequences

Overall percentages of correct gender identifications in CV sequences are extremely high: 98.25 % for French listeners and 98.5 % for American English listeners. Nonetheless, mean degree of certainty slightly varies from one stimulus to another. Therefore, correlation tests were conducted between this variable and several acoustic parameters: vowel formant frequencies (F1, F2 and F3), H1-H2 value in open vowels, mean f0 and duration. Results for the tests performed on French listeners' data are presented in Table 3.

These tests provide interesting information about French listeners' strategies. First of all, it appears that vowel formant frequencies probably played an important role in such gender identification task. Several significant correlations are found between this acoustic parameter and listeners' main degrees of certainty. They are more frequently found with F2 values (4 significant correlations) than with F1 and F3 values (1 significant correlation for each of them): this suggests that F2 frequency

might have a larger influence on French listeners' judgments than F1 and F3 frequencies.

Table 3 Results of Pearson correlation tests conducted on French listeners' degrees of certainty in CV sequences and various acoustic parameters (only significant correlations are shown in the table).

Pearson correlations in C+V sequences for French listeners	**Y variable :** mean degree of certainty
F2 (Hz) vowel [i] – Male voices	$r(16) = -0.569$; $z = -2.331$; $p < 0.02$*
F2 (Hz) vowel [a] – Male voices	$r(16) = -0.491$; $z = -1.899$; $p < 0.05$*
F1 (Hz) vowel [u] – Female voices	$r(16) = 0.634$; $z = 2.695$; $p < 0.01$*
F2 (Hz) vowel [u] – Female voices	$r(16) = 0.728$; $z = 3.331$; $p < 0.001$*
F3 (Hz) vowel [u] – Female voices	$r(16) = 0.759$; $z = 3.582$; $p < 0.001$*
F2 (Hz) vowel [u] – Male voices	$r(16) = -0.497$; $z = -1.927$; $p < 0.05$*
Mean f0 (Hz) all stimuli – Male voices	$r(48) = -0.426$; $z = -3.053$; $p < 0.01$*

Concerning the other acoustic parameters, it appears that stimulus duration probably did not influence French listeners in CV sequences. Likewise, no clear tendency was observed for H1-H2 values, which provides an indication of phonation type: this parameter might be irrelevant for French listeners in identifying gender from voice. On the other hand, it seems that mean fundamental frequency influenced listeners' judgments, at least in male voices: the lower mean f0, the higher the degrees of certainty for this type of items.

Similar correlation tests were performed on American English listeners' data. Results of these tests are shown in Table 4.

Table 4 Results of Pearson correlation tests conducted on American English listeners' degrees of certainty in CV sequences and various acoustic parameters (only significant correlations are shown in the table).

Pearson correlations in C+V sequences for American English listeners	**Y variable :** mean degree of certainty
F2 (Hz) vowel [u:] – Female voices	$r(16) = 0.649$; $z = 2.788$; $p < 0.01$*
H1-H2 (dB) vowel [æ] – Male voices	$r(16) = -0.497$; $z = -1.967$; $p < 0.05$*
Mean f0 (Hz) all stimuli – Male voices	$r(48) = -0.341$; $z = -2.38$; $p < 0.02$*

Contrary to French listeners, Pearson correlations are not conclusive with vowel formant frequencies: only one positive correlation was found in vowel [u:]'s F2 for female voices. This suggests that formant frequencies have less influence on American English listeners than on French listeners in a gender identification from voice task. Moreover, it appears that stimulus duration does not play a significant role in CV sequences for American English listeners, which is consistent with what was found for French listeners.

Two other acoustic parameters seem to have strongly influenced American English listeners' judgments. Unlike French listeners, H1-H2 value is significantly correlated to the degree of confidence expressed on male voices: the lower H1-H2 is (i.e. the smaller glottal open quotient is), the more confident American English listeners are. An opposite tendency was found for female voices, but the correlation does not reach significance. These facts suggest that phonation type is a relevant clue in identifying gender for American English listeners. Likewise, mean fundamental frequency also seems to somewhat influence the listeners' degree of certainty. Indeed, a strong negative correlation was found in male voices.

2.3 Entire CVCV words

For CVCV words, overall percent correct identification scores are again very high: 99.93 % for French listeners and 99.41 % for American English listeners. Thus, correlation tests cannot be performed on this variable. However, as we did with CV sequences, Pearson correlation tests were conducted between mean degree of certainty and several acoustic parameters: H1-H2 value in open vowels, mean f0, duration, but also f0 range. Results for French listeners' data are presented in Table 5.

Table 5 Results of Pearson correlation tests conducted on French listeners' degrees of certainty in dissyllabic words and various acoustic parameters (only significant correlations are shown in the table).

Pearson correlations in dissyllabic words for French listeners	**Y variable :** mean degree of certainty
Mean f0 (Hz) all stimuli – Female voices	$r(54) = 0.273$; $z = 1.997$; $p < 0.05$*
Duration (ms) all stimuli – Female voices	$r(54) = 0.478$; $z = 3.712$; $p < 0.001$*

Results support what was previously found with CV sequences. First of all, H1-H2, which provides an estimation of phonation type, did not

seem to influence French listeners in identifying gender from dissyllabic words. However, a significant positive correlation was found between stimulus duration and mean degree of certainty in female voices only. This suggests that a slow speech rate might be associated to female voices in French listeners.

Mean fundamental frequency seems to have influenced French listener's judgments on female voiced. The higher mean f0, the more French listeners tended to be sure of their choice. Nevertheless, the opposite correlation was not found with male voices. Thus, the effect of mean f0 on French listeners was probably quite reduced. It is highly probable that resonant frequencies and particularly vowel formant frequencies, which were not tested for this category of stimuli, strongly influenced French listeners' judgments as in CV sequences.

Finally, an interesting tendency was found concerning f0 range. Despite their non-significance, we can notice that there is a positive correlation with mean degree of certainty in female voices, and a negative correlation in male voices. Further research in that direction might suggest that the more f0 range is wide in female voices or narrow in male voices, the more French listeners tend to be confident in their choice.

Similar correlation tests were conducted on American English listeners' data. The results are shown in Table 6.

Table 6 Results of Pearson correlation tests conducted on American English listeners' degrees of certainty in dissyllabic words and various acoustic parameters (only significant correlations are shown in the table).

Pearson correlations in dissyllabic words for American English listeners	**Y variable :** mean degree of certainty
H1-H2 (dB) in vowel [æ] – Male voices	$r(18) = -0.531$; $z = -2.291$; $p < 0.05$*
Mean f0 (Hz) all stimuli – Female voices	$r(54) = 0.587$; $z = 4.812$; $p < 0.0001$*
Duration (ms) all stimuli – Female voices	$r(54) = 0.523$; $z = 4.143$; $p < 0.0001$*
Duration (ms) all stimuli – Male voices	$r(54) = -0.379$; $z = -2.849$; $p < 0.01$*

For American English listeners, tendencies observed in CV sequences are also broadly confirmed in dissyllabic words. An interesting tendency appears when male and female voices are considered separately. One observes a strong and significant positive correlation between word

duration and mean degree of certainty for female voices, and a significant negative correlation between these two variables for male voices. This suggests that for American English listeners a slow speech rate would be associated with female voices, and a fast speech rate to male voices.

Results also suggest that mean fundamental frequency played a crucial role for American English listeners: the higher mean f0 was, the more listeners were sure of their choices. An opposite tendency was found in male voices, but does not reach significance. Contrary to French listeners, H1-H2 difference seems to have great influence on American English listeners: the lower H1-H2 was (i.e. the lower glottal open quotient was), the more listeners were sure of their judgments on male voices.

3. Conclusion – Discussion

The gender identification from speech experiment described in the present paper differs from previous studies on several aspects. First, it was conducted jointly on French and American English listeners, using stimuli in respective languages. Secondly, the gating paradigm was used. Finally, stimuli were natural, not resynthesized speech. Thanks to these methodological peculiarities it has been possible to confirm previous results, but also to get new findings about gender identification form speech in human listeners.

In voiceless initial consonants, mean percent correct gender identification reaches 66 % in French listeners and 76 % in American English listeners, way above chance level. These results indicate that listeners are capable to identify gender from very short speech segments. These results are quite similar to those obtained by Whiteside (1998b). When the consonant was voiced, the listeners' scores widely and significantly increased in both languages. However, one can find here some evidence of cross-language variation. For American English listeners, percentage of correct gender identification reaches 99.33 %, versus a mere 94.5 % for French listeners.

In isolated initial vowels, American listeners' scores are similar to those obtained in voiced consonants. On the other hand, they are slightly increasing for French listeners, either in terms of percent correct identification or of degree of certainty. This might indicate that vowel formants would be an important acoustic parameter in identifying gender for French listeners, but not necessarily for American English listeners.

In CV sequences and dissyllabic words, very high percentages of correct identification and degrees of certainty were obtained in both languages. Correlations between mean degrees of certainty and several

acoustic parameters have broadly confirmed the above mentioned tendencies: mean fundamental frequency appeared to have greater influence on American English listeners than on French listeners, while vowel formant frequencies seemed to be more relevant for French listeners.

It is therefore possible to state that the apparently contradictory results found in studies led either on American English listeners or on French listeners are actually compatible. Indeed, it is highly probable that mean f0 is the most salient acoustic clue for American English listeners, as suggested in Coleman (1976), Lass, & al. (1976), and Pausewang Gelfer & Mikos (2005), and that resonant frequencies and particularly vowel formant frequencies are the most important acoustic parameter for French listeners, as mentioned in Arnold (2012). The weaker role played by vowel formants for American English listeners could be accounted for by the large regional variations existing on this parameter in American English. Another possible explanation is that the formants of English vowels are less relied on since they are known to be less stationary than those in French (Pike, 1947).

Besides these two main acoustic parameters, it appeared that phonation type, through H1-H2 measurements, was also a relevant clue for gender identification from speech, but only for American English speakers. The stronger H1 was relatively to H2 (i.e. high GOQ, thus breathy voice quality), the more listeners were sure of their "female speaker" categorization. The weaker H1 was, the more the listeners were sure of their "male speaker" categorization. Cross-gender acoustic differences on this parameter in American English speakers was described by several authors (Klatt & Klatt, 1990)d, but no experimental study had proved the perceptual salience of this parameter in gender identification from speech yet.

An interesting, though not significant, tendency was found for f0 range. French listeners' results suggest that a large f0 range would help the identification of female voices, while a narrow f0 range would help the identification of male voices. No similar tendency was observed for American English listeners, which supports previous findings from Pépiot (2011). This latter study showed that intonation could be a relevant clue in gender identification from speech for French listeners but not for American English listeners.

In dissyllabic words, it appeared that speech rate could be a clue for identifying gender from voice. In female voices, the longer the words were, the more American English and French listeners tended to be confident in their judgments. An opposite correlation was found in male voices, but only for American English speakers. Just as phonation type, cross-gender

acoustic differences on this parameter have already been described (Byrd, 1994), but no experimental study was conducted to investigate its possible influence on listeners in a gender identification from speech task.

Regarding the results of the present experiment and taking into account previous similar studies, we drew up a summary chart (Table 7). This table mentions an estimated degree of influence for each acoustic parameter involved in gender identification from dissyllabic words, depending on the listeners' first language (French or American English).

Table 7 Estimated degree of influence of different acoustic parameters in gender identification from voice in French and American English listeners. The symbol "ø" means no influence, and a scale goes from "*" (very weak influence) to "***" (very strong influence).**

GENDER IDENTIFICATION FROM VOICE IN DISSYLLABIC WORDS		
Acoustic parameter	**Estimated degree of influence**	
	French listeners	*American English listeners*
Mean f0	****	*****
Vowel formant frequencies	*****	***
Phonation type	ø	***
Speech rate	*	**
F0 range	**	ø
Consonant noise frequencies	*	*

These data have to be interpreted with caution. Indeed, it is known that the importance of acoustic parameters in gender identification from speech is relative. For instance, when a speaker displays a mean f0 located at an ambiguous level, the influence of other acoustic parameters are immediately reinforced (Coleman, 1976; Pépiot, 2011, 2014a). Moreover, the current study presents certain limits. First of all, only 4 different speakers for each language (2 women and 2 men) were recorded to make up the stimuli.

Nevertheless, considering the important cross-language variations we found between French and American English listeners through the present study, it seems very likely that listeners' strategies for identifying gender from speech are language-dependent. This suggests that listeners have devised socially-constructed mental representations of what "female voices" and "male voices" are.

These findings could be relevant to transgender people. Since voice and speech are parts of gender identity, having a good knowledge of which acoustic parameter one has to work on could help these persons to be

effectively perceived as a woman or a man. Beyond mean fundamental frequency, which is often spontaneously adapted by transgender people (Spencer, 1988), other features such as vowel formant frequencies, phonation type, speech rate or f0 range could be improved. A list of relevant acoustic parameters should ideally be drawn up for each language, regarding the relative influence they have in gender identification from speech. Data presented in Table 7 could then be used for American English and Parisian French languages.

REFERENCES

André & al. 2003: André, C., Ghio, A., Cavé, C., & Teston, B**.** Perceval: a computer-driven system for experimentation on auditory and visual perception. // *Proceedings of the 15th International Congress of Phonetic Sciences, Barcelona*, *3-9 August 2003*. Barcelona: UAB, 2003, 1421-1424.

Arnold 2012: Arnold, A. Le rôle de la fréquence fondamentale et des fréquences de résonance dans la perception du genre. // *TIPA – Travaux Interdisciplinaires sur la Parole et le Langage*, 2012, Nº 28, 1-18.

Byrd 1994: Byrd, D. Relations of sex and dialect to reduction. // *Speech Communication*, 1994, Nº 15, 39-54.

Coleman 1976: Coleman, R. O. A comparison of the contributions of two voice quality characteristics to the perception of maleness and femaleness in the voice. // *Journal of Speech and Hearing Research*, 1976, Nº 19, 168-180.

Grosjean 1980: Grosjean, F. Spoken word recognition processes and the gating paradigm. // *Perception and Psychophysics*, 1980, 28, 267-283.

Johnson 2005: Johnson, K. Speaker normalization in speech perception. // Pisoni, D. & Remez, R. *The Handbook of Speech Perception*. Oxford: Blackwell Publishers, 2005, 363-389.

Klatt & Klatt 1990: Klatt, D. H., & Klatt, L. C. Analysis, synthesis and perception of voice quality variations among female and male talkers. // *Journal of the Acoustic Society of America*, 1990, Nº 87, 820-857.

Lass & al. 1976: Lass, N. J., Hughes, K. R., Bowyer, M. D., Waters, L. T., & Bourne, V. T. Speaker sex identification from voiced, whispered, and filtered isolated vowels. // *Journal of the Acoustic Society of America*, 1976, Nº 59, 675-678.

Pausewang Gelfer & Mikos 2005: Pausewang Gelfer, M., & Mikos, V. The relative contributions of speaking fundamental frequency and formant frequencies to gender identification based on isolated vowels. // *Journal of Voice*, 2005, Nº 19, 544-554.

Pépiot 2011: Pépiot, E. Voix de femmes, voix d'hommes : à propos de l'identification du genre par la voix chez des auditeurs anglophones et francophones. // *Scientific Works – Philology,* Paisii Hilendarski University of Plovdiv – Bulgaria, 2011, Nº 49, 418-430.

Pépiot 2014a: Pépiot, E. Voix et genre: un état de la question. // *La langue, la Voix, la Parole*. Ibrahim, A. H. (éd.). Paris: CRL, 2014, 53-86.

Pépiot 2014b: Pépiot, E. Male and female speech: a study of mean f0, f0 range, phonation type and speech rate in Parisian French and American English speakers. // *Social and Linguistic Speech Prosody.* Proceedings of the 7th International Conference on Speech Prosody – Dublin, 2014, 305-309.

Pépiot 2016a: Pépiot, E. Female and male speech: a study of vowel formants and consonant noise in Parisian French and American English speakers. // *Scientific Works – Philology,* Paisii Hilendarski University of Plovdiv – Bulgaria, 2016, Vol. 53, 610-621.

Pépiot 2016b: Pépiot, E. 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. // *Actes de la conférence conjointe JEP-TALN-RECITAL, Jul 2016, Paris, volume 1 : JEP, 2016,* 759-767, <https://jep-taln2016.limsi.fr/actes/-Actes%20JTR-2016/V01-JEP.pdf> (2 June, 2018).

Pike 1947: Pike, K. L. On the phonemic status of English diphthongs. // *Language*, 1947, Nº 23, 151-159.

Schwartz 1968: Schwartz, M. F. Identification of speaker sex from isolated voiceless fricatives. // *Journal of the Acoustical Society of America*, 1968, Nº 43, 1178-1179.

Spencer 1988: Spencer, L. Speech characteristics of male-to-female transsexuals: A perceptual and acoustic study. // *Folia Phoniatrica et Logopaedica*, 1988, Nº 40, 31-42.

Whiteside 1998a: Whiteside, S. P. Identification of a speaker's sex: a study of vowels. // *Perceptual and Motor Skills*, 1998, Nº 86, 579-584.

Whiteside 1998b: Whiteside, S. P. Identification of a speaker's sex: a fricative study. // *Perceptual and Motor Skills*, 1998, Nº 86, 587-591.