



HAL
open science

Aerodynamic and durational cues of phonological voicing in whisper

Yohann Meynadier, Yulia Gaydina

► **To cite this version:**

Yohann Meynadier, Yulia Gaydina. Aerodynamic and durational cues of phonological voicing in whisper. Interspeech, Aug 2013, Lyon, France. pp.335-339. hal-01211117

HAL Id: hal-01211117

<https://hal.science/hal-01211117>

Submitted on 5 Oct 2015

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Aerodynamic and durational cues of phonological voicing in whisper

Yohann Meynadier, Yulia Gaydina

Aix-Marseille Université, CNRS UMR 7309, LPL, Aix-en-Provence, France

yohann.meynadier@lpl-aix.fr, gaydina_yulia@live.fr

Abstract

This paper presents analyses on the phonological voicing contrast in whispered speech, which is characterized by the absence of vocal fold vibrations. In modal speech, besides glottal vibration, the contrast between voiced and unvoiced consonants is realized by other phonetic correlates: e.g. consonant and pre-consonantal vowel durations, intraoral pressure differences. The analysis of these voicing cues shows they are well preserved in whispered speech. The results are consistent with previous studies supporting the perception of the voicing contrast even without physical voicing. Additionally, the patterns observed for P_o during obstruents could reflect constraints connected to competing needs of intelligibility and of discretion in whisper.

Index Terms: phonetics, voicing, whisper, aerodynamics, segmental duration

1. Introduction

Whisper is a mode of phonation naturally used in order to reduce the perceptibility of speech. In modal phonation (“normal voice”), the fully adducted vocal cords allow them to vibrate for voiced segments. In whisper, the vocal folds are adducted only in their front part, with a small posterior opening between arytenoids left for exhaust air. This Y-like configuration allows the establishment of glottal aerodynamic turbulences generating the noisy sound source.

Besides the lack of voicing and the different laryngeal configuration, whispered voice induces many changes in different phonetic dimensions: (i) spectral changes as a significant loss of energy especially at low frequencies, a flattening of high frequencies and higher vowel formants [1-7]; (ii) temporal changes as a lengthening of syllables or segments or a lowering of speech rate [3, 5, 6, 8, 9]; (iii) aerodynamic changes as an increase of airflow and air consumption [10-12] and equivocal findings on subglottal and intraoral pressures, [11, 12] for a review; (iv) supraglottal articulatory changes towards some kind of hyperarticulation [13-15].

Concerning the perception of whispered speech, studies on a small number of languages reported that most segmental and suprasegmental information is more or less preserved, as (i) vowel identity [16-18]; (ii) consonant place and manner [19]; even (iii) intonation, accent [3, 6, 20-22] or tone [23] (but see [24, 25]); and (iv) voicing feature in [6, 19, 26-31], that is targeted in our study.

This study focuses on the duration of pre-consonantal vowels and, on the duration and intraoral pressure of obstruents as secondary phonetic properties involved in maintaining the perception of phonological voicing in whispered speech, i.e. without phonetic (physiological and acoustic) voicing. In modal speech, apart from the periodic vibration of the vocal folds, these properties (or their acoustic consequences) are part of numerous secondary phonetic cues commonly reported for voicing [32-34]. The duration of

consonants and pre-consonant vowels has long been observed: (i) vowels are longer before voiced than before voiceless consonants and, (ii) voiceless obstruents are longer than voiced ones [2], see [35, 36] for a review.

While the voiced-voiceless difference in duration of pre-consonantal vowels is not still well understood [2, 37, 38], the difference of obstruent duration has usually been explained by the *Aerodynamic Voicing Constraint* [39]. A minimal difference between subglottal (P_s) and intraoral (P_o) pressures ($\Delta P = P_s - P_o$) is required to start and to maintain vocal vibration. Indeed, below 1 to 2 hPa [33, 39] vibration ceases because of the weakness of translaryngeal airflow. This constraint at work for voiced obstruents would account for their frequent devoicing and their shortening to avoid devoicing. Voiceless consonants are not subject to this aerophonatory requirement. Also, the laryngeal resistance of adducted vocal folds during voicing (*Glottal Resistance Hypothesis*) induces a decrease of the translaryngeal airflow. Hence, the voiced obstruents show a lower P_o than the unvoiced ones [39], see [40] for a short review. The lower P_o results in a lower energy of the burst or the friction noise of voiced obstruents, which is a cue of voicing.

With regard to whisper, one may logically think that these aerodynamic constraints (on consonant duration and spectral energy) do not operate due to a static abduction of the vocal folds during speech, as largely described in phonetic textbooks, see [41] for references. But, some other phoneticians suggested (without empirical evidences) a more dynamical view of the glottis during whisper [33, 42] (see [30] for other references) that could support the voicing contrast. A narrower glottis (as in Y-like whisper phonation) would distinguish the whispered voiced obstruents from the unvoiced obstruents phonated with a larger glottis close to the V-like voiceless abduction produced in modal speech. To our knowledge, only three studies ([30] on English, [43] on Moroccan Arabic, [44] on French) have attempted to test this proposition by direct empirical data on the state of the glottis during whispered speech. While aerodynamics provides indirect information on the configuration of the glottis [45], few studies have used indirect P_o measurement to assess glottal adjustments related to the consonant voicing contrast in whisper. Unfortunately, they reported equivocal findings. In Japanese, higher P_o for whispered unvoiced than voiced obstruents are established by Higashiwaka [28] but not by Soda et al. [46]. In English, Murry and Brown [47] did not find any support to this contrast. Weismer and Longstreth’s [41] results showed only a trend. However, their measurements of airflow peak and P_o slope seem to attest a possible narrower glottis adjustment for whispered voiced obstruents, even if this seems insufficient to cause such a significant large P_o drop as in modal speech. So, at this time, direct (optical) or indirect (P_o) observations provides no clear and strong evidence that the *Glottal Resistance Hypothesis* is operating in whispered speech.

Concerning the duration of vowels and consonants in whisper, more data are available. First, works on modal speech

show that, besides VOT cue (especially for English), gating experiments on the closing-phase consonant duration ([48] on French) and on the pre-consonantal vowel duration ([49, 50] on English) modify significantly the perception of the voicing feature of obstruents. Although no such psycholinguistic test has been done in whispered speech, some studies show that the perception of the voicing contrast is generally preserved [19, 26-31]. Only a pilot test by Vercherand [6] supports these results for French. None of them allows connecting precisely this performance in whisper on any identified secondary properties of voicing. Nevertheless, vowel and consonant durations are always approached as major secondary cues of consonant voicing. The reason is that studies on the production of vowel and consonant durations in whisper all converge to account for a preserved durational contrast for vowels and obstruents, even reduced, in whispered speech, as in English [29, 30, 51, 52], Serbian [5], Dutch [53], Russian and Hungarian [31] and French [6].

Our study on whispered speech production in French focuses on these questions. In a first experiment, we examine durational cues of the voicing feature, in order to confirm and to complete previous results obtained by Vercherand [6]. The second experiment tests the *Glottal Resistance Hypothesis* by means of indirect P_0 measurement. This is the first attempt to support Malécot and Peebles's direct observation [44], concerning glottal adjustments as function as the phonological voicing of French obstruents in whispered speech.

2. Voicing-dependant durations

This first experiment concerns the analysis of the acoustic duration of voiced vs. unvoiced obstruents and of their pre-consonantal vowels in modal vs. whispered speech.

2.1. Corpus and data

Recorded in an anechoic room, 4 French speakers (2 males, 2 females, not linguistic experts or students) read aloud at normal speaking rate two randomized lists of words presented in isolation. After a brief training session, they produced each list 5 times alternatively for each phonation mode.

A first list was composed of 12 non-sense words pooled with 36 fillers. The targeted words patterned as VC_1VC_2V , wherein V was always /e/ and, C_1 and C_2 obstruent target always different from the voiced-unvoiced pairs: /b-p/, /t-d/, /k-g/, /f-v/, /s-z/ and /ʃ-ʒ/, e.g. /epeʒe/ (written *épéjé*) and /edese/ (written *édécé*). Fillers met the same pattern, but included at least one non-target consonant, e.g. /ekene/ (written *équéné*) or /eleme/ (written *éléme*). The occurrence frequency of consonants was balanced inside the list. Only consonants in C_1 position are analyzed here.

A second list consisted of 12 target conjugated verbs, randomly pooled with 72 fillers of the same phonological template. The obstruents were intervocalic and in (unstressed) non-final syllable of words. (Syllable-initial obstruents in word-final position were recorded too, but not analyzed here). The preceding vowel was /e/ in all cases, except 4 (/a/). To control for the following vowel, voiced-unvoiced consonants are part of minimal pairs, e.g. (*il*) *écoutait* /ekute/ [(he) listened to] vs. (*il*) *égouttait* /egute/ [(he) dripped]. Speakers read aloud only the words between brackets to produce only three-syllable sequences.

The duration of pre-consonantal vowels and consonants were acoustically measured with Praat.

2.2. Measures and analyses

To ensure same measurements between phonation modes, segmentation was mostly done on spectrogram, from the F2-F3 onset and offset of vowels. The beginning and end of vowels correspond respectively to the end and beginning of consonants. The manual labeling was carried out with Praat.

For consonants, two articulatory phases were identified (Figure 1). For stops, the closure (*T*) was labeled from the end of the preceding vowel to the beginning of the burst. The release (*R*) was measured from the beginning of the burst to the beginning of the following vowel. The friction noise after the burst may have formant traces, the release phase can also be seen as CV transition phase. For fricatives, the steady-state of the constriction (*T*) was separated from its release (*R*) identified at the beginning of stronger formant traces of the following vowel during fricative noise (i.e. CV transition).

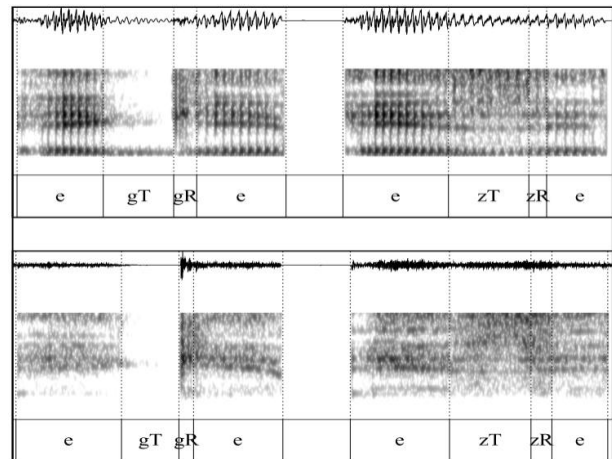


Figure 1: Segmentation of consonant and preceding /e/ for [ege(pe)] (left) and [eze(te)] in modal (top) and whisper (bottom) speech.

The durations of consonant and preceding vowels were analyzed by repeated-measure ANOVA tests with *speaker* and *lexical status* (word vs. non-sense word) as random factors. The independent factors are *phonation* mode (modal vs. whisper), *voicing* (voiced vs. unvoiced) and *manner* of articulation (stop vs. fricative). Lexical and non-sense words were pooled. All reported values are mean durations (in ms).

2.3. Results and discussion

2.3.1. Obstruents

First, obstruent durations significantly, but weakly, increase in whisper (107 ms) compared to modal phonation (98 ms) [$F(1,3) = 13.881$; $p = 0.03368$]. This does not completely fit Vercherand's results [6] on text-reading French where a phonation effect is only for 2/4 speakers on unvoiced fricatives. This is in agreement with studies on others languages [5, 8, 9] related to slower speech rate in whisper.

Concerning the voicing contrast, as expected in modal phonation, unvoiced obstruents (118 ms) are longer than the voiced ones (78 ms) [$F(1,3) = 161.79$; $p = 0.00105$]. Unvoiced fricatives and stops are longer than voiced counterparts: respectively 133 and 102 ms vs. 85 and 74 ms, as showed in Figure 2. Similarly, in whispered speech, unvoiced consonants are significantly 31 ms longer than voiced ones

$[F(1,3) = 56.014; p = 0.00494]$: respectively, 136 vs. 99 ms for fricatives and 108 vs. 82 ms for stops. The durational difference associated with the phonological voicing contrast of obstruents is also kept in whisper. The difference between unvoiced and voiced fricatives reduces from modal ($\Delta = 48$ ms) to whisper ($\Delta = 37$ ms). For stops, the difference remains constant: 28 in modal speech and 26 ms in whisper.

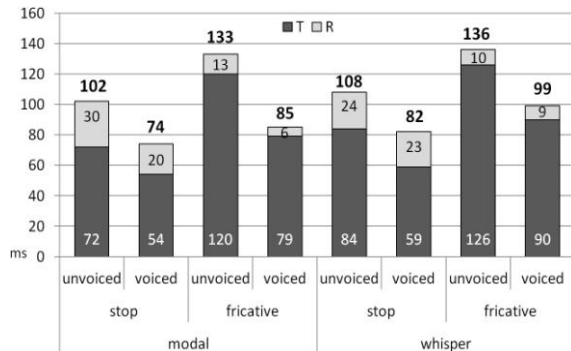


Figure 2: Duration of consonants (top), of their steady-state (bottom) and of their release (middle)

The analysis brings out a noticeable difference between the steady-state and the release of obstruents (Figure 2). In modal words, both these phases take significantly part to the durational distinction between voiced and unvoiced consonants. It is not the case in whisper where the release duration seems to play no role in the voicing contrast.

For whispered stops, the closure is significantly longer for the unvoiced (84 ms) than voiced (59 ms) ones $[F(1,3) = 55.134; p = 0.00505]$. The difference ($\Delta = 25$ ms, i.e. 84 to 59 ms from voiced to unvoiced obstruents) is enhanced compared to that in modal speech ($\Delta = 18$ ms, 72 to 54 ms). No effect of voicing on the release duration of whispered stops is observed (24 and 23 ms for unvoiced and voiced ones), unlike modal stops produced with a 50% difference (30 and 20 ms for unvoiced and voiced ones) $[F(1,3) = 13.670; p = 0.03434]$.

Whispered voiced fricatives present shorter durations of the steady-state phase voiced (90 ms) $[F(1,3) = 49.841; p = 0.00584]$. The difference is 36 ms with unvoiced fricatives (126 ms). Unlike stops, there is a slight reduction compared to 41 ms measured in modal phonation. Here again, no effect is observed on the duration of their release in whisper, in opposition to modal phonation.

Thus, it is mainly the steady-state of obstruents which seems to support voicing information in whisper, and not the release duration in this French corpus. Few data are available to discuss this point here. Languages without aspiration reported closure and global duration lengthening of obstruents without any remark on the release duration [5, 6, 31, 53]. In languages with aspiration (English), consonant [29], closure [15] or aspiration [52] are only alternatively reported. Kinsey [52] observed an effect on release and not on closure.

Finally, our main results on longer values for unvoiced than voiced obstruents in whisper confirm and complete previous data.

2.3.2. Vowels

A longer vowel precedes voiced obstruents and vice versa (Figure 3). Similar significant differences are observed whatever the phonation mode and the manner of articulation of

obstruents: $\Delta = 15$ ms in modal speech $[F(1,3) = 93.103; p = 0.00236]$ and $\Delta = 16$ ms in whisper $[F(1,3) = 245.76; p = 0.00056]$.

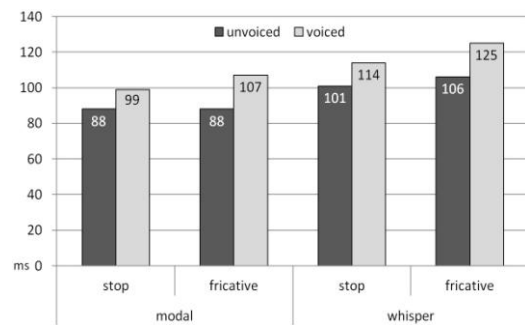


Figure 3: Duration of pre-consonantal vowels

Our data strengthen results observed once on French [6]. They corroborate that the phonological voicing of obstruents is also marked in French by a difference of the pre-consonantal vowel duration in whisper, as in modal speech, just as in others languages [29, 31, 51-53].

3. Glottal Resistance Hypothesis

This second experiment focuses on the difference of intraoral pressure (P_o) as function of voiced/unvoiced contrast of labial obstruents in whispered speech. The P_o is seen as an indirect measurement of the glottal opening in order to assess the *Glottal Resistance Hypothesis* in whisper.

3.1. Corpus and data

In an anechoic room, 8 French speakers (6 females, 2 males) read aloud at normal speaking rate a balanced randomized list of 72 isolated words, once in modal voice and once in whisper. The labial obstruents /p-b/ and /f-v/ were produced in initial and final positions of monosyllabic words, and in median position of bisyllabic words. Vocalic contexts were: (i) /a ε o / for initial and final position, e.g. *vache* /vaʃ/ (cow), *beige* /beɜ/ (beige), *pomme* /pɔm/ (apple) or, *chef* /ʃɛf/ (chief), *rap* /ʁap/ (rap), *robe* /ʁɔb/ (dress); (ii) /a_ a ε o / for median position, e.g. *savate* /savat/ (old shoe), *affaire* /afɛʁ/ (case), *rapport* /ʁapɔʁ/ (report). Each vowel context was iterated in two different words, totaling per speaker 18 occurrences of each consonant in each phonation mode.

Acoustic signal, oral airflow and P_o were recorded synchronously with the EVA aerophonometer, [54] for technical details. P_o was measured via the mouth in oral cavity by a special fine tube which ends approximately 1 cm behind the central incisors. When labial closure or labiodental constriction was realized, a positive increase of P_o is captured.

3.2. Measures and analyses

P_o measures were carried out with Phonedit (<http://www.lpl-aix.fr/~lpldev/phonedit>). Expressed in hectopascal (hPa), the maximal P_o value was measured at the peak of the P_o curve reached during the consonant.

The statistical analyzes were repeated-measure ANOVA. *Speaker*, *syllable position* and *vowel context* are random factors. *Phonation* (modal vs. whisper), *voicing* (voiced vs. unvoiced) and *manner* of articulation (plosive vs. fricative) are independent factors. All reported values are mean P_o .

3.3. Results and discussion

P_o does not differ regularly between modal and whispered phonation. The variability between subjects across phonations shows that P_o is not systematically lower in whisper (Table 1). It is confirmed by previous equivocal findings on the decrease P_s or P_o in whisper [11, 12, 28, 41, 46, 47].

Table 1. P_o peak of labial obstruents in modal (Mod) and whisper (Whi) phonations. Speakers in rows.

	/p/		/b/		/f/		/v/	
	Mod	Whi	Mod	Whi	Mod	Whi	Mod	Whi
F1	4.36	5.02	2.6	4.46	3.95	4.35	2.42	3.27
F2	5.27	3.89	2.72	3.43	4.47	3.53	1.96	2.87
F3	8.00	6.08	5.21	5.56	7.42	6.92	6.43	5.23
F4	6.26	7.49	3.94	6.18	6.4	7.43	4.73	4.99
F5	4.73	4.11	4.08	3.34	5.69	4.68	4.34	3.85
F6	4.59	4.42	3.05	4.32	4.59	3.86	2.25	3.64
M1	5.31	4.46	2.72	3.58	5.27	4.21	3.48	3.26
M2	5.29	4.51	2.74	3.35	5.13	3.63	4.41	3.09
mean	5.48	5	3.38	4.28	5.37	4.83	3.75	3.78

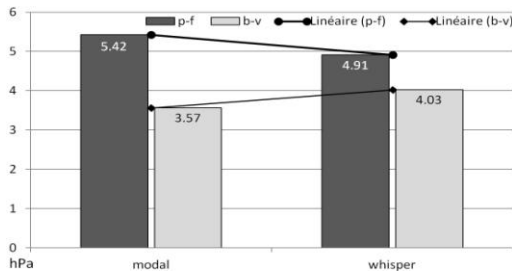


Figure 4: P_o peak of labial obstruents

The data clearly show that *voicing* has a main effect in modal speech [$F(1, 7) = 135$; $p = 0.000008$] and, by contrast with previous studies, strikingly still in whisper [$F(1, 7) = 26.2023$; $p = 0.001371$]. Among all factors and interactions, only *voicing* [$F(1, 7) = 138.162$; $p = 0.00007$] and the *voicing*phonation* interaction [$F(1,7) = 16.631$; $p = 0.0047$] account for a significant effect (Figure 4).

In both phonations, the unvoiced labial obstruents were produced with a greater P_o than voiced ones: respectively 5.42 vs. 3.56 hPa in modal voice and, 4.91 vs. 4.02 hPa in whisper. This result supports similar trends observed by Weismer and Longstreth in English [41] and Higashiwaka [28] in Japanese. In our knowledge, these results are unique for French.

This finding could be seen as a new support for the *Glottal Resistance Hypothesis* in whispered speech. Indeed, a dynamic adjustment of the glottis as function of the phonological voicing of consonant could explain in part the P_o differences observed here. As expected [33, 42], among others (see [30]), the glottal opening would reduce from unvoiced to voiced consonants in whisper. According to these authors, it is produced by a narrower glottis exclusively for whispered voiced consonants. This is partly supported by Mills' endoscopy study [30] who finds no glottal difference for unvoiced obstruents in function of phonations and, a trend to a closer glottis for voiced compared to unvoiced obstruents. Malécot and Peebles [44] did a single similar observation on French, but not Zeroual et al. [43] for Moroccan. Nonetheless, the *voicing*phonation* interaction (Figure 4, lines for interaction) may indicate a more complex pattern than a manifestation of a voicing contrast in whisper.

First, the P_o differences supporting the voicing contrast is significantly halved from modal (1.86 hPa) to whispered (0.89

hPa) phonation. Thus, if the contrast is maintained at the glottal level in whisper, it is also weakened. Secondly, the reduction of the contrast is due to opposite behaviors of voiced and unvoiced consonants. On the one hand, the P_o increases for voiced obstruents (from 3.47 to 4.04 hPa) from modal to whispered phonation. On the other, P_o decreases for unvoiced ones (from 5.42 to 4.91 hPa). Indeed, the *Glottal Resistance Hypothesis* could only account for the behavior of voiced consonants produced by a Y-like whisper configuration [33, 42]: P_o reduces, because of the narrower glottis inducing a greater glottal resistance against the translaryngeal airflow. For unvoiced ones, it would suggest no change in P_o from modal to whispered phonation, because of the constant V-like abduction, expected for their nil (silent) phonation. Hence, why should P_o decrease for unvoiced obstruents?

A first hypothesis is that P_s is lower in whisper. Nevertheless there are no previous results which would allow us to draw this conclusion, see [11] vs. [12]. Additionally, this could no longer account for increasing P_o of voiced obstruents. Another possibility relies on the perceptual need to low acoustic energy required to whisper. Narrowing the glottis, and so strengthening the glottal resistance against the translaryngeal airflow, P_o decreases. P_o drop is a way to low the turbulent noise produced at the supraglottal constriction. This can be related to the -4dB weak difference in noise intensity of fricatives between modal and whispered speech [5]. A weak decrease of intensity could be due to a weak difference of glottal abduction. The glottal constriction of whispered voiceless obstruents reported by Zeroual et al. [43] would support this hypothesis.

So, P_o patterns in whisper may result from conflicting perceptual constraints relative to intelligibility or phonological requests (as voicing contrast) and needs for discreet communication by soft quiet speech.

4. Conclusions and perspectives

First, this study supports that the production of the voicing contrast in whispered speech is marked by secondary cues for vowel and obstruent durations, close to those commonly attested for modal speech. For French, only the previous work by Vercherand [6] approached this question.

Moreover, the more original contribution of the paper concerns the analysis of intraoral pressure (P_o) of obstruents in function of their phonological voicing in whisper compared to modal speech. The observed patterns could surface constraints connected to competing needs for intelligibility and discretion in whisper. So, while the P_o increases between modal and whispered phonations for voiced obstruents, it decreases for unvoiced ones. The former pattern could be based on a phonological voicing control, the latter on a constraint of low-intensity speech mode. In both cases, perception seems to play a central role. But, it is no clear what role these kinds of physiological constraints could also have.

Voicing contrast in whisper seems to be another new paradigm of interest to weigh the different constraints of various kinds and levels (physiological, phonological, perceptual or communicative) that operate on speech and communication processes.

5. Acknowledgements

This work is granted by the Ministère de la Recherche et l'Enseignement supérieur and the Région PACA.

6. References

- [1] Schwartz, M. F., "Power spectral density measurements of oral and whispered speech," *J. Speech and Hear. Res.*, 13: 445-446, 1970.
- [2] Lehiste, I., *Suprasegmentals*, MIT Press, 1970.
- [3] Faraco, M., *Comparaison des intonations affirmative et interrogative en voix normale et chuchotée*. PhD, Université de Provence, 1984.
- [4] Ito, T., Takeda, K. and Itakura, F., "Analysis and recognition of whispered speech", *Speech Com.*, 45(2): 139-152, 2005.
- [5] Jovicic, S.T. and Saric, Z., "Acoustic analysis of consonants in whispered speech", *J. of Voice*, 22(3): 263-274, 2008.
- [6] Vercherand, G., *Production et perception de la parole chuchotée en français: analyse segmentale et prosodique*, PhD, Université de Paris 7, 2010.
- [7] Sharifzadeh H. R., McLoughlin, I. V. and Russell, M. J., "A comprehensive vowel space for whispered speech". *J. Voice*, 26(2), e49-e56, 2012.
- [8] Schwartz, M. F., "Syllable duration in oral and whispered reading", *J. Acou. Soc. of Am.*, 41(5): 1367-1369, 1967.
- [9] Schwartz, M. F., "Bilabial closure durations for /p/, /b/, and /m/ in voiced and whispered vowel environments," *J. Acou. Soc. of Am.*, 51(6): 2025-2029, 1972.
- [10] Schwartz, M. F., "Air consumption, per syllable, in oral and whispered speech". *J. Acou. Soc. of Am.*, 43: 1448-1449, 1971.
- [11] Stathopoulos, E. T., Hoit, J. D., Hoit, T. J., Watson, P. J. and Solomon, N. P., "Respiratory and laryngeal function during whispering", *J. Speech and Hear. Res.*, 34: 761-767, 1991.
- [12] Konnai, R. M., *Whisper and phonation: aerodynamic comparisons across adduction and loudness levels*, PhD, Bowling Green State University, 2012.
- [13] Higashikawa, M., Green, J. R., Moore, C. A. and Minifie, F. D., "Lip kinematics for /p/ and /b/ production during whispered and voiced speech", *Folia Phon. et Logo.*, 55(1): 17-27, 2003.
- [14] Traunmüller, H., and Ericksson, A., "Acoustic effects of variation in vocal effort by men, women, and children", *J. Acou. Soc. of Am.*, 107(6): 3438-3451, 2000.
- [15] Osfar, M. J. O., "Articulation of whispered alveolar consonants", Master of Arts, University of Illinois, 2011.
- [16] Kallail, K.J., and Emanuel, F. W., "The identifiability of isolated whispered and phonated vowel samples", *J. Phonetics*, 13: 11-17, 1985.
- [17] Tartter, V. C., "Identifiability of vowels and speakers from whispered syllables", *Attention, Perception, & Psychophysics*, 49(4): 365-372, 1991.
- [18] Eklund, I., and Traunmüller, H., "Comparative study of male and female whispered and phonated versions of the long vowels of Swedish", *Phonetica*, 54(1): 1-21, 1997.
- [19] Tartter, V. C., "What's in a whisper?", *J. Acou. Soc. of Am.*, 86(5): 1678-1683, 1989.
- [20] Trim, J., "Cues to the recognition of some linguistic features of whispered speech in English". 6th ICPhS Proc., 919-923, 1970.
- [21] Nicholson, H. and Teig, A.H., "How to tell beans from farmers: cues to the perception of pitch accent in whispered Norwegian", *Nordlyd*, 31(2): 315-325, 2003.
- [22] Heeren, W. and Heuven van, V., "Perception and production of boundary tones in whispered Dutch", 10th Interspeech Proc., 2411-2414, 2009.
- [23] Jensen, M.C., "Recognition of word tones in whispered speech", *Word*, 14: 186-196, 1958.
- [24] Miller, J. D., "Word tone recognition in Vietnamese whispered speech", *Word*, 17: 11-15, 1961.
- [25] Gao M., *Tones in whispered Chinese: articulatory features and perceptual cues*. Master of Arts, University of Victoria, 2002.
- [26] Dannenbring, G. L., "Perceptual discrimination of whispered phoneme pairs", *Perceptual and Motor Skills*, 51: 979-985, 1980.
- [27] Munro, M J., "Perception of 'voicing' in whispered stops". *Phonetica*, 47(3-4): 173-181, 1990.
- [28] Higashikawa, M., "Perceptual, acoustical and aerodynamic study of whispering". *Nippon Jibiinkoka Gakkai Kaiho*, 97: 1268-1280, 1994. (*Paper in Japanese cited from the English abstract*)
- [29] Mills, T.I.P., *Cues to voicing contrasts in whispered Scottish obstruents*, Master of Science, University of Edinburgh, 2003.
- [30] Mills, T.I.P., *Speech motor control variables in the production of voicing contrasts and emphatic accent*, PhD, University of Edinburgh, 2009.
- [31] Knyazev, S., "O svyazi osobennostej artikulatsionnoj bazy govora s harakterom protivopostavleniya gluhih i zvonkih, tverdyh i myagkih soglasnyh", *Sovremennye russkie govory*, 1991.
- [32] Delattre, P., "Les indices acoustiques de la parole", *Phonetica*, 2: 108-118 & 226-251, 1958.
- [33] Catford, J. C., *Fundamental problems in phonetics*, Edinburgh University Press, 1977.
- [34] Lisker, L., "Voicing in English: a catalogue of acoustic features signaling /b/ versus /p/ in trochees", *Language and Speech*, 29(1): 3-11, 1986.
- [35] Abdell-Beruh, N. B., "The stop voicing contrast in French sentences: contextual sensitivity of vowel duration, closure duration, voice onset time, stop release and closure voicing", *Phonetica*, 61: 201-219, 2004.
- [36] Fowler, C. A., "Vowel duration and closure duration in voiced and unvoiced stops: there are no contrast effects here", *J. Phonetics*, 20: 143-165, 1992.
- [37] Lisker, L., "On 'explaining' vowel duration variation", *Glossa*, 8: 233-246, 1974.
- [38] Raphael, L. J., "The physiological control of durational differences between vowels preceding voiced and voiceless consonants in English", *J. Phonetics*, 3(1): 25-34, 1975.
- [39] Ohala, J. J., "Aerodynamics of phonology", 4th Seoul Internat. Conf. on Linguistics Proc., 92-97, 1997.
- [40] Malécot, A., "An experimental study of force of articulation", *Studia Linguistica*, 9: 35-44, 1955.
- [41] Weismer, G. and Longstreth, D., "Segmental gestures at the laryngeal level in whispered speech: evidence from an aerodynamic study," *J. Speech and Hear. Res.*, 23: 383-392, 1980.
- [42] Laver, J., *Principles of phonetics*. Cambridge University Press, 1994.
- [43] Zeroual, C., Esling, J. H. and L. Crevier-Buchman, "Physiological study of whispered speech in Moroccan Arabic", 6th Interspeech Proc., 1069-1072, 2005.
- [44] Malécot, A. and Peebles, K., "An optical device for recording glottal adduction-abduction during normal speech", *Zeitschrift für Phonetik*, 18: 545-550, 1965.
- [45] Ladefoged, P., Maddieson, I., and Jackson, M., "Investigating phonation types in different languages", in O. Fujimura [Ed], *Vocal fold physiology*, 297-317, Raven Press, 1988.
- [46] Soda, T., Nishida, Y. and Suwoya, H., "Intraoral pressure change in Japanese consonants", *Otologica Fukuoka*, 13(1): 34-43, 1967. (*cited in [47]*)
- [47] Murry, T. and Brown, W. S., "Peak intraoral pressures in whispered stop consonants", *J. Phonetics*, 4: 183-187, 1976.
- [48] Wajskop, M. and Sweerts, J., "Voicing cues in oral stop consonants", *J. Phonetics*, 1: 121-130, 1973.
- [49] Raphael, L.J., "Preceding vowel duration as a cue to the perception of the voicing characteristic of word-final consonants in American English", *J. Acou. Soc. of Am.*, 51: 1296-1303, 1972.
- [50] Hogan, J. T. and Rozsypal, A. J., "Evaluation of vowel duration as a cue for the voicing distinction in the following word-final consonant", *J. Acou. Soc. of Am.*, 67(5): 1764, 1980.
- [51] Sharf, D. J., "Vowel duration in whispered and in normal speech", *Language and Speech*, 7: 89-97, 1964.
- [52] Kinsey, R. H. "Phonological voice in whispered speech," *Linguistics 110 Paper*, 1-15, 2005.
- [53] Velde van der, D. J. and Heuven van, V., "Compensatory strategies for voicing of initial and medial stops and fricatives in whispered speech in Dutch." 17th ICPhS Proc., 2058-2061, 2011.
- [54] Ghio, A. and Teston, B., "Evaluation of the acoustic and aerodynamic constraints of a pneumotachograph for speech and voice studies", *Internat. Conf. on Voice Physiology and Biomechanics Proc.*, 55-58, 2004.