Aerodynamic tool for phonology of voicing

Yohann Meynadier

AERODYNAMIC TOOL FOR PHONOLOGY OF VOICING

Yohann Meynadier

Aix Marseille Université, CNRS, LPL UMR 7309, 13100, Aix-en-Provence, France
yohann.meynadier[at]lpl-aix.fr

ABSTRACT

The paper addresses the question of the phonetic implementation of phonological voicing in French. It is investigated by means of aerodynamic measures made at the subglottal, glottal and supraglottal stages of the production of fricatives in modal voice and in soft and loud whisper for one speaker. The results show that the [±voice] feature is systematically associated only with the categorical constrictions of the glottis, even without vocal fold vibration.

Keywords: voicing, whisper, subglottal pressure, intraoral pressure, glottis, vocal effort, French

1. INTRODUCTION

The voicing contrast meets different phonological feature specifications across languages and theories, e.g. voice, fortis/lenis, tense/lax, aspiration, spread glottis, glottal constriction, see e.g. [8, 12]. The question of their association with identifiable phonetic properties is also largely discussed. Our study deals with some points relative to this large question in French. How is the [±voice] feature phonetically implemented? Is it associated with subglottal, glottal and/or supraglottal components? Is voicing defined by vocal fold vibration?

Aerodynamics provides a powerful tool to study synchronously all components of the speech production system and to test phonological proposals by empirical experiments, e.g. [15, 27]. Aerodynamic studies globally report that intraoral pressure (*Pio*) in various languages reflects the [±voice] features of obstruantes, inter alia [1, 3, 6, 7, 18, 21, 25, 30], but see e.g. [17]. As *Pio* must be lower than the subglottal pressure (*Psg*) to reach the mechanical vibration of the folds, *Pio* is viewed as an essential control parameter of voicing [28].

But *Pio* is a somewhat ambiguous phonetic correlate of voicing, because its value depends on various mechanisms taking place at different levels of the speech production system: (i) at a subglottal level as function of the expiratory effort; (ii) at a glottal level by the laryngeal configuration; (iii) at a surpraglottal level by various articulatory mechanisms [28]. Its ambiguity results in different proposals concerning the physical level of the phonetic implementation of the voicing contrast in production. For example, based on laryngectomized speech, Kobayashi [14] thus suggests that *Pio* is phoneme specific as function of voicing, but not the glottis configuration itself. Conversely, based on *Pio* during whispered speech in French, Meynadier and Gaydina [21] concluded that *Pio* depends mainly on the glottal configuration directly associated with the voicing feature, as Netsell [25] and Slis [30].

To confirm the last proposal, in this study, the aerodynamic production of the voicing contrast in whisper is focused on French. Whispered speech is characterised by an absence of vocal fold adduction, hence of their periodic vibration. Despite this lack, the few studies which have been made report that the perception of voicing contrasts in various languages is quite well preserved in whisper [5, 11, 22, 23, 32], French included [9, 20, 33]. Concerning the voicing production in whisper, the *Pio* results of the few existing studies are equivocal (see §4).

To clear the *Pio* ambiguity and to assess which manoeuvres at which level of speech production are responsible for the voicing contrast, we record aerodynamic parameters at the successive stages of the air pathway in modal and whispered speech. Therefore the subglottal expiratory force is estimated by *Psg*, the size of the laryngeal opening by the glottal resistance to lung airflow, the vocal tract adjustments by *Pio* and the size articulatory constriction by the supraglottal resistance to the phonatory airflow. The fricatives are focused here, because simultaneous aerodynamic measures at all speech production stages are possible.

2. METHOD

2.1. Corpus

One French speaker (the author) read at normal speaking rate a randomized list of 6 isolated non-sense words in modal (‘voiced phonation’), whispered (‘voiceless phonation at the ear of a listener’) and loud whispered (‘voiceless phonation to a 2 meter remote listener’) voices. Each list was alternatively read 5 times in each phonation mode.

The 6 fricatives /f-v/, /s-z/ and /ʃ-ʒ/ are the target phonemes inside non-sense words with the pattern /e$C_{[\alpha\ \mathrm{voice}]}eC_{[-\alpha\ \mathrm{voice}]}$e/. The acoustic segmentation of consonants was processed on the oscillogram only. The beginning and the end of the fricatives was located at points of change in the signal waveform during the transition towards and from the vowel.

2.2. Aerodynamics

The acoustic signal, the oral airflow (*Oaf*), the subglottal (*Psg*) and supraglottal (*Pio*) pressures were recorded synchronously with the EVA station [10]. *Oaf* was captured by a silicon mask covering the mouth. *Pio* was recorded with a catheter (internal diameter: 2 mm) introduced through the nasal tract to just below the uvula. *Psg* was measured via a tracheal catheter (i.d.: 1.3 mm) punctured between the cricoïd cartilage and the first tracheal ring. Aerodynamic measurements (in l/s and hPa) were carried out with the free Phonedit software [29].

The airflow and air pressure *y* values were automatically extracted at the peak of *Pio* during the fricative, i.e. the *Pio* maximum and, *Psg* and *Oaf* *y* values at this *x* time point. The *y* values were normalized by subtracting the minimum *y* value reached after the end of each word, that is at the *x* time point of the first zero-crossing of *Psg*. The post-calibration allows us to exclude the local up-steps, down-steps or declination of the signals.

From these measures, two aerodynamic parameters (in $hPa/l_{/S}$) were computed: (1) the resistance of the glottis to lung airflow; (2) the resistance of the vocal tract constriction to airflow.

$$(1)\quad Rglo = \frac{Psg - Pio}{Oaf} \qquad (2)\quad Rsup = \frac{Pio - Patm}{Oaf}$$

where *Patm* equals to 0 as a fixed constant of atmospheric pressure.

The lack of glottal vibration of the whispered fricatives was checked by the EGG signal, synchronously recorded with the aerodynamic data.

2.3. Statistics

The statistical analyses were simple linear regressions, two-way ANOVAs and post-hoc Fisher tests on these aerodynamic measures and parameters (N = 3 consonants * 2 positions * 2 voicing * 3 phonation conditions * 5 repetitions = 178; one /f/ and one /v/ excluded).

The C_1 or C_2 position, the place of articulation and the repetitions are random factors. *Phonation* (modal vs. whisper vs. loud whisper), and *Voicing* ([+voice] vs. [-voice]) are the independent factors (significant levels: *<.05; **<.01; ***<.0001).

3. RESULTS

Simple linear regressions were preliminarily carried out to assess the redundancy of measures and parameters.

The aerodynamic measures (*Psg*, *Pio* and *Oaf*), used to compute the two parameters of airflow resistance, appears largely independent of each other. Significant R^2 values (p<.05 at least) vary from .055 to .289 among the 3 phonation conditions. The high correlation observed between *Pio* and *Psg* (.496, pooled conditions) essentially reflects the categorical change in pressure due to the vocal effort increasing from whisper to modal voice to loud whisper. In split conditions, it's lost.

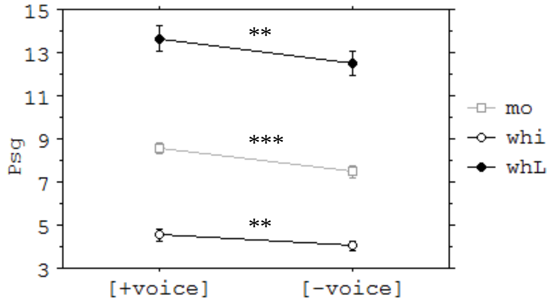
Only poor correlations between the aerodynamic parameters (*Psg*, *Rglo* and *Rsup*) emerge: $.079 \leq R^2 \geq .423$. The higher correlation between *Psg* and *Rglo* seems here to capture a categorical effect of the consonant voicing. When the regression is computed separately for voiced (R^2=.197; p=.014) and voiceless (R^2=.183; p=.0185) fricatives, it disappears. It could be the sign of the main similar impact of the adduction/abduction of the vocal folds on the amplitude of *Psg* and of *Rglo*.

So, the linear regression analyses show that only few partial and weak correlations link the aerodynamic parameters. When some correlations exsit, they reflect in fact more categorical effects relative to the different conditions of phonation, the vocal effort or the consonant voicing. It is rather clear that these aerodynamic parameters are quite independent. Therefore, they capture specific aerodynamic mechanisms involved in the different stages of the production of fricatives: subglottal (*Psg*) vs. glottal (*Rglo*) vs. supraglottal (*Rsup*).

3.1. Vocal effort (*Psg*)

The ANOVAs report a main significant effect of *Phonation* [F(2,172)=1060.52; p<.0001] and of *Voicing* [F(1,172)=33.655; p<.0001] on *Psg*, but no significant *Phonation*Voicing* interaction.

Figure 1: *Psg* as function of *Phonation*/*Voicing* (confident interval of 95% in bars; * for significant levels for *Voicing*; idem in the next figures)



Psg increases strongly (by around 4-5 hPa step) from whisper to modal voice to loud whisper. That is as function of the expiratory force needed to produce the expected acoustic energy in each speech type.

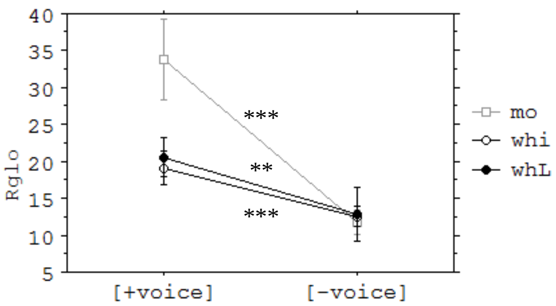
Comparatively, the significant difference due to *Voicing* is limited to 0.5-1 hPa, as already measured in French [6, 21] or in English [25], and simulated by Ohala [26]. But, it is systematic over the 3 conditions of phonation. With or without vocal fold vibration, *Psg* is higher for [+voice] than [-voice] fricatives. This fine difference may be a mechanical and passive consequence of a stronger resistance of the closer glottis for [+voice] fricatives. Hence, even if the glottis remains open in whisper, a narrower glottal constriction for /v z ʒ/ could explain the weak but constant increase of *Psg*.

Here, the expiratory force doesn't seem to be controlled by the phonological voicing specification of the segment, but to be a means of controlling the global level of the acoustic energy of the voice, as function of phonation modes and planned intensity.

3.2. Glottal resistance (*Rglo*)

Phonation [F(2,172)=12.018; p<.0001] and interaction *Phonation*Voicing* [F(2,172)=15.805; p<.0001] have significant effects on *Rglo*. But, Figure 2 shows clearly that the effects are only due to the specific adduction of vocal folds used to produce the vibration source of [+voice] fricatives in modal voice. Otherwise, the glottal resistance remains similar in whisper whatever the vocal effort. In loud whisper, it isn't mainly supported or controlled by the laryngeal system as glottal tension or configuration, but by the expiratory effort (*Psg*).

Figure 2: *Rglo* as function of *Phonation*/*Voicing*



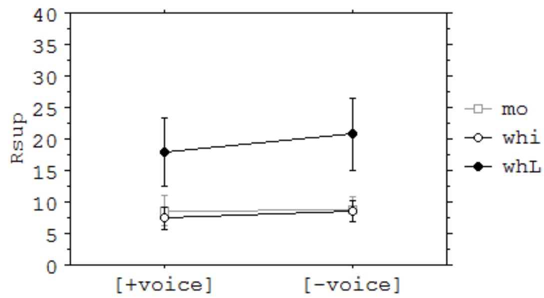
Conversely, the glottal resistance significantly varies as function of *Voicing* [F(1,172)=92.323; p<.0001]. In modal voice as in whisper, the [+voice] fricatives are produced with a significantly greater *Rglo*. While the difference (22.2 hPa/l_s) is three times in modal voice than in whisper the [+voice] fricatives are realised with higher *Rglo* of around 7 hPa/l_s than the [-voice] fricatives. Moreover, only 3 categorical levels of glottal resistance emerge: [-voice] < whispered [+voice] < modal [+voice].

As a quite similar *Psg* was observed as function of *Voicing* (§3.1), the aerodynamic resistance of the glottis may decrease with its degree of opening (and/or of the laryngeal constriction). Therefore, the [-voice] specification for voiceless fricatives seems to be phonetically produced with the same glottis opening, whatever the phonation mode and the vocal effort. But mostly, even if the vibration of the vocal folds (by their adduction) is prohibited, as in whispered speech, the phonological [+voice] specification of fricatives seems to lead to a glottal constriction, for both vocal efforts produced. So, this result confirms that the phonological voicing is physiologically produced by different degrees of constriction of the glottis or of the laryngeal sphincter.

3.3. Supraglottal resistance (*Rsup*)

A significant effect on *Rsup* is only reported for *Phonation* [F(2,172)=25.414; p<.0001]. The fricatives produced in loud whisper show a greater resistance of articulatory constriction to airflow. This could result from the highest *Psg* (§3.1) plus a quite open glottis (§3.2).

Figure 3: *Rsup* as function of *Phonation*/*Voicing*



Concerning *Voicing*, no substantial differences can be observed, despite a tenuous increase of *Rsup* from [+voice] to [-voice] fricatives for all phonation conditions. So, the results may support an absence of a specific control of the [±voice] specification at the articulatory constriction.

3.4. *Pio* and *Oaf* as unstable echoes of voicing

Phonation and *Voicing* have significant effects on *Pio* [F(2,172)=95.169; p<.0001 and F(1,172)= 11.405; p=.0009], and on *Oaf* [F(2,172)=68.86; p<.0001 and F(1,172)= 22.71; p<.0001]. An interaction is reported only for *Oaf* due to the large increase of airflow from the [+voice] to [-voice] fricatives only in modal voice.

Concerning the differences between [+voice] and [-voice] fricatives, *Pio* (Figures 4 and 6) and *Oaf* (Figures 5 and 6) seem to be mainly echoes of *Rglo*. When Δ*Rglo* is high, as in modal voice, *Oaf* and *Pio* reflect the voicing contrast. When Δ*Rglo* is low, as in whisper, these echoes are very weak and/or variable. *Pio* and *Oaf* may be seen as passive responses to pneumo-phonatory configurations. This interpretation is also supported by the parallel differences of *Pio* and *Psg* across the 3 modes of phonation (Figure 6). As a matter of fact, *Pio* and *Oaf* are the less consistent aerodynamic correlates of the voicing contrast across the vocal conditions.

Figure 4: *Pio* as function of *Phonation*/*Voicing*

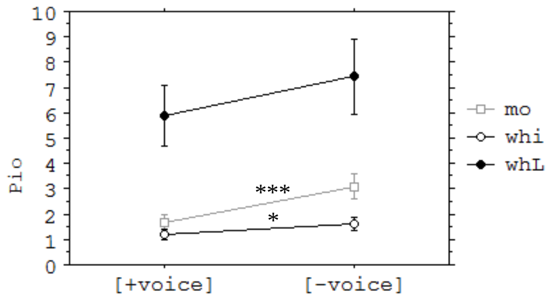


Figure 5: *Oaf* as function of *Phonation*/*Voicing*

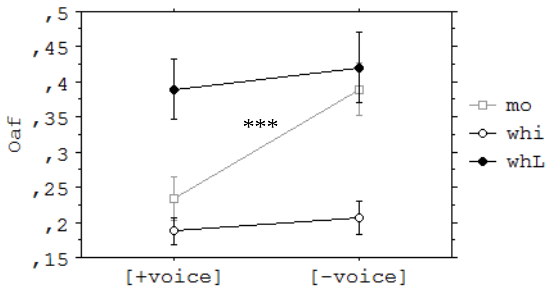


Figure 6: Significant differences and Δ values (below) between [+voice] (*Fv*) and [-voice] (*Fs*) fricatives at each aerodynamic levels.

air pathway		**mod**	**whi**	**whL**
speech	*Oaf*	Fs > Fv -0.16	ns (-0.02)	ns (-0.03)
vocal tract	***Rsup***	**ns** (-0.2)	**ns** (-1.2)	**ns** (-2.9)
	Pio	Fs > Fv -1.4	Fs > Fv -0.4	ns (-1.5)
glottis	***Rglo***	**Fs < Fv** 22.2	**Fs < Fv** 6.6	**Fs < Fv** 7.7
lungs	*Psg*	Fs < Fv 1.1	Fs < Fv 0.5	Fs < Fv 1.1

4. DISCUSSION

The summary of the phonological voicing effects on fricatives at the successive stages of speech production (Figure 6) shows that the relationships of aerodynamic parameters along the airway confirm that the implementation of the [±voice] feature states only at the glottis level. It does not seem to take place above or below the larynx in phonation with or without vocal fold vibration.

First, as the air pressure control above the vocal folds is essential for voicing, the well-known *Aerodynamic Voicing Constraint* [28] may also be involved in whisper, in French [21]. Many mechanisms inside the vocal tract can support or strengthen it, see e.g. [14, 28]. A main supraglottal control factor of *Pio* is the size of the articulatory constriction for which the supraglottal resistance to phonatory airflow (*Rsup*) can be an indirect measure. In our study, non-discriminatory *Rsup* data rules out the use of the articulatory constriction as a systematic mechanism to support the voicing contrast. The *Pio* value itself is also a result of other kinds of active or passive supraglottal manoeuvres or adjustments according to voicing. But here again, no consistent variation of *Pio* data allow us to support such systematic mechanisms in whisper. *Pio* seems only a variable echo of the glottal and subglottal components. It is confirmed by numerous equivocal findings on *Pio* values as function of voicing in whisper in Japanese, [11] vs. [31], or in English, [13] and [25] vs. [24] vs. [34].

The air pressure control below the vocal folds is also essential for voicing. In our study, the small differences of *Psg* may be understood as only passive mechanical echoes of the degree of glottal opening. This is supported by previous studies on English [25] and on Dutch [30] in modal voice and in whisper. So, the expiratory effort doesn't seem to be actively adjusted phoneme by phoneme.

Finally, the alternative proposal considers that the [±voice] contrast is essentially controlled by the states of the glottis alone, e.g. [2, 16], see [22] for a review. The resistance to expiratory airflow is an aerodynamic index of the glottis size. Our *Rglo* data show clearly that a stronger resistance of the glottis is systematically produced for [+voice] even in whisper, and whatever the pneumo-phonatory effort. Conversely, [-voice] is always realised with non-distinctive *Rglo* values, hence a stable wide glottal configuration. Even without complete adduction and vibration of the vocal folds, as in whispered speech, the glottis is closer for [+voice] fricatives and wider for [-voice] ones. These results are supported by the similar aerodynamic studies on whisper by Netsell [25] in English and by Slis [30] in Dutch. They are also consistent with Malécot and Peebles' single electrophotoglottographic observation [20], and the only direct videofibroscopic studies (that I know of) in French by Crevier-Buchman et al. [5] and in English by Mills [22] on whispered speech.

As an outcome on the phonology of French, the [+voice] feature may be associated with a narrow constriction of the glottis, and/or the laryngeal sphincter, to produce the acoustic source, periodic or noisy. Its phonetic correlate, such as the vocal fold vibration, may only be determined by the mode of phonation and not the phonological representation.

5. REFERENCES

[1] Butcher, A. 1992. Intraoral pressure as an independent parameter in oral stop contrasts. *Proc. 4th Speech Sc. Tech.* Brisban 286–291.

[2] Catford, J.C. 1977. *Fundamental problems in phonetics*. Edinburgh: Edinburgh University Press.

[3] Collier, R., Lisker, L., Hirose, H., Ushijima, T. 1979. Voicing in intervocalic stops and fricatives in Dutch. *J. Phonetics.* 7, 357–373.

[4] Crevier-Buchman, L., Vaissière, J., Henrich, N., Vincent, C., Hans, S., Brasnu, D. 2009. Laryngeal behaviour in whispered voice: a study using high speed imaging. *The Voice Foundation's 38th Annual Symposium: Care of the Professional Voice*. Philadelphia.

[5] Dannenbring, G.L. 1980. Perceptual discrimination of whispered phoneme pairs. *Perceptual and Motor Skills* 51, 979–985.

[6] Demolin, D., Giovanni, A., Hassid, S., Heim, C., Lecuit, V., Soquet, A. 1997. Direct and indirect measurements of subglottic pressure. *Proc. Larynx 1997* Marseille, 69–72.

[7] Fischer-Jørgensen, E. 1972. PTK et BDG français en position intervocalique accentuée. In: Valdman, A. (ed.), *Papers in linguistics and phonetics to the memory of Pierre Delattre*. Mouton, 143–200.

[8] Fuchs, S. 2005. Articulatory correlates of the voicing contrast in alveolar obstruent production in German. *ZAS Papers in Linguistics* 41.

[9] Fux, T. 2012. *Vers un système indiquant la distance d'un locuteur par transformation de sa voix.* Université de Grenoble: Thèse doctorale.

[10] Ghio, A., Teston, B. 2004. Evaluation of the acoustic and aerodynamic constraints of a pneumotachograph for speech and voice studies. *Proc. 4th Int. Conf. on Voice Physiology and Biomechanics* Marseille, 55–58.

[11] Higashikawa, M. 1994. Perceptual, acoustical and aerodynamic study of whispering. *Nippon Jibiinkoka Gakkai Kaiho* 97, 1268–1280. (abstract)

[12] Keating, P. A. 1984. Phonetic and phonological representation of stop consonant voicing. *Language* 60(2), 286–319.

[13] Klich, R.J. 1982. Effects of speech level and vowel context on intraoral air pressure in vocal and whispered speech. *Folia Phoniatrica et Logopaedica* 34(1), 33–40.

[14] Kobayashi, N. 1987. Intraoral air pressure variations for consonantal voicing distinction in speech without larynges. *An. Bull. RILP* 21, 193–201.

[15] Ladefoged, P., Maddieson, I., Jackson, M. 1988. Investigating phonation types in different languages. In: Fujimura, O. (ed.), *Vocal fold physiology*. Raven Press, 297–317.

[16] Laver, J. 1994. *Principles of phonetics*. Cambridge: Cambridge University Press.

[17] Lisker, L. 1970. Supraglottal Air Pressure in the Production of English Stops. *Language and Speech* 13(4), 215–230

[18] Malécot, A. 1955. An experimental study of force of articulation. *Studia Linguistica* 9, 35–44.

[19] Malécot, A., Peebles, K. 1965. An optical device for recording glottal adduction-abduction during normal speech. *Zeitschrift für Phonetik* 18, 545–550.

[20] Meynadier, Y., Dufour, S., Gaydina, Y. 2013. Duration as perceptual voicing cue in whisper. *Abstract booklet 6th PaPI* Lisbon.

[21] Meynadier, Y., Gaydina, Y. 2013. Aerodynamic and durational cues of phonological voicing in whisper. *Proc.14th Interspeech* Lyon, 335–339.

[22] Mills, T.I.P. 2009. *Speech motor control variables in the production of voicing contrasts and emphatic accent*. University of Edinburgh: Ph Dissertation.

[23] Munro, M.J. 1990. Perception of voicing in whispered stops. *Phonetica* 47(3-4), 173–181.

[24] Murry, T., Brown, W. S. 1976. Peak intraoral pressures in whispered stop consonants. *J. Phonetics* 4, 183–187.

[25] Netsell, R. 1969. Subglottal and intraoral air pressures during the intervocalic contrast of /t/ and /d/. *Phonetica* 20(2-4), 68–73.

[26] Ohala, J. J. 1975. A mathematical model of speech aerodynamics. In Fant, G. (ed.), *Speech Communication*, Almqvist & Wiksell, vol. 2, 65–72.

[27] Ohala, J.J. 1997. Aerodynamics of phonology. *Proc. 4th Int. Conf. on Linguistics* Seoul, 92–97.

[28] Ohala, J.J. 2011. Accommodation to the aerodynamic voicing constraint and its phonological relevance. *Proc. 17th ICPhS* Hong-Kong, 64–67.

[29] Phonedit free software http://www.lpl-aix.fr/~lpldev/phonedit.

[30] Slis, I.H. 1970. Articulatory measurements on voiced, voiceless and nasal consonants. *Phonetica* 21(4), 193–210.

[31] Soda, T., Nishida, Y., Suwoya, H. 1967. Intraoral pressure change in Japanese consonants. *Otologica Fukuoka* 13(1), 34–43.

[32] Tartter, V.C. 1989. What's in a whisper? *J. Acou. Soc. of Am.* 86(5), 1678–1683.

[33] Vercherand, G. 2010. *Production et perception de la parole chuchotée en français: analyse segmentale et prosodique*. Université de Paris 7 : Thèse doctorale.

[34] Weismer, G., Longstreth, D. 1980. Segmental gestures at the laryngeal level in whispered speech: evidence from an aerodynamic study. *J. Speech and Hear. Res.* 23, 383–392.