

Duration as perceptual voicing cues in whisper

Yohann Meynadier, Sophie Dufour & Yulia Gaydina

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

yohann.meynadier[at]lpl-aix.fr, sophie.dufour[at]lpl-aix.fr, gaydina_yulia[at]live.fr

This study concerns the production and the perception of the phonological voicing contrast in whispered speech in French.

Whisper is a mode of phonation naturally used in order to reduce the perceptibility of speech, mainly substituting the periodic sound source of modal voice by a noisy sound source. Whispered voice induces many changes: (i) intensity lowering, frequency flattening and formant rising [1-6]; (ii) lengthening of speech units and speech rate decrease [3, 5-7]; (iii) increase of airflow and air consumption [8]; and (iv) some kind of hyperarticulation [9]. Concerning perception, segmental and suprasegmental information is generally well perceived, with a recognition level above the chance: (i) vowel identity [10]; (ii) consonant place and manner [11]; even (iii) intonation, accent [3, 6, 12] or tone [13]; and strikingly (iv) voicing feature [6, 11, 14-16], as it is targeted here for French.

This study focuses on duration of pre-consonantal vowels and obstruents as secondary phonetic cues in production and perception of the phonological voicing in whispered speech, i.e. without phonetic (physiological and acoustic) voicing. In modal speech, these properties are part of numerous secondary phonetic cues commonly reported for voicing [17]. Duration of consonants and pre-consonant vowels are long frequently observed: (i) vowels are longer before voiced than voiceless consonants [2, 19, 20] and, (ii) voiceless obstruents are longer than voiced ones [2, 21, 22] (for a review and discussion).

A first experiment on production confirms that the phonological voicing contrast is also realized in whisper. Alternatively in modal and whisper phonations, 4 French speakers read 12 non-sense and 12 lexical words embedded the voiced and unvoiced obstruents /b-p/, /t-d/, /k-g/, /f-v/, /s-z/ and /ʃ-ʒ/ in word-median position. The list-reading recordings were experimentally controlled: random order, fillers, anechoic room, etc.

As in modal phonation, in whisper acoustic durations show that unvoiced consonants are significantly 31 ms longer than voiced ones. 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. Similar significant differences are observed whatever the phonation mode for pre-consonantal vowels: $\Delta = 11$ ms before stops and $\Delta = 19$ ms before fricatives. So, the durational differences associated with the phonological voicing contrast of obstruents are also kept in whisper production.

In a second experiment on perception, durations of consonant-median closure and pre-consonantal vowel were acoustically manipulated to fit the duration of the counterpart member of a minimal pair (e.g. [d] to [t]... and vice versa). The proportion of the temporal lengthening or shortening of segments were based on the empirical results of the production test. The perception test was experimentally controlled: stimuli, random order, fillers, intensity level, experimental materials, etc.

First analyses show that the perception decreases slightly for whispered voiced obstruents (close to 90% of correct responses), but surprisingly very dramatically for unvoiced ones (around the chance level). Crucially, the results showed that consonant duration has more impact on the recognition of the voicing than vowel duration. These effects are cumulative, depending on the case. These results are discussed in relation to previous studies.

Finally, to our knowledge, this study is the first attempt (at least in French) to clear duration effects on voicing perception in whisper.

References

- [1] Schwartz, M. F., "Power spectral density measurements of oral and whispered speech," *J. of 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] Jovicic, S.T. and Saric, Z., "Acoustic analysis of consonants in whispered speech", *J. of Voice*, 22(3): 263-274, 2008.
- [5] Vercherand, G., *Production et perception de la parole chuchotée en français: analyse segmentale et prosodique*, PhD, Université de Paris 7, 2010.
- [6] Sharifzadeh H. R., McLoughlin, I. V. and Russell, M. J., "A comprehensive vowel space for whispered speech". *J. of Voice*, 26(2), e49-e56, 2012.
- [7] Schwartz, M. F., "Syllable duration in oral and whispered reading", *J. of the Acou. Soc. of Am.*, 41(5): 1367-1369, 1967.
- [8] Schwartz, M. F., "Air consumption, per syllable, in oral and whispered speech". *J. of the Acou. Soc. of Am.*, 43: 1448-1449, 1971.
- [9] Osfar, M. J. O., "Articulation of whispered alveolar consonants," Master dissertation manuscript, University of Illinois, 2011.
- [10] Tartter, V. C., "Identifiability of vowels and speakers from whispered syllables," *Attention, Perception, & Psychophysics*, 49(4): 365–372, 1991.
- [11] Tartter, V. C., "What's in a whisper?", *J. of the Acou. Soc. of Am.*, 86(5): 1678–1683, 1989.
- [12] Heeren, W. and Heuven van, V., "Perception and production of boundary tones in whispered Dutch", *10th Interspeech Proc.*, 2411-2414, 2009.
- [13] Jensen, M.C., "Recognition of word tones in whispered speech", *Word*, 14: 186-196, 1958.
- [14] Dannenbring, G. L., "Perceptual discrimination of whispered phoneme pairs", *Perceptual and Motor Skills*, 51: 979-985, 1980.
- [15] Munro, M J., "Perception of 'voicing' in whispered stops". *Phonetica*, 47(3-4): 173-181, 1990.
- [16] Mills, T.I.P., *Speech motor control variables in the production of voicing contrasts and emphatic accent*, PhD, University of Edinburgh, 2009.
- [17] Lisker, L., "Voicing in English: A catalogue of acoustic features signaling /b/ versus /p/ in trochees", *Language and Speech*, 29(1): 3-11, 1986.
- [18] Lisker, L., "On 'explaining' vowel duration variation", *Glossa*, 8: 233-246, 1974.
- [19] Raphael, L. J., "The physiological control of durational differences between vowels preceding voiced and voiceless consonants in English", *J. of Phonetics*, 3(1): 25–34, 1975.
- [20] 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.
- [21] Fowler, C. A., "Vowel duration and closure duration in voiced and unvoiced stops: there are no contrast effects here", *J. of Phonetics*, 20: 143-165, 1992.