Vowel duration and tongue root advancement: Results from an exploratory study of the relation between voicing and vowel duration

Stefano Coretta

02/05/2018

1 Introduction

This paper reports a previously undocumented correlation between vowel duration and degree of tongue root advancement. In an exploratory study of the articulatory correlates of stop voicing, it has been found that tongue root advancement—a known mechanism that facilitates voicing during stop closure—can be implemented not only during the closure of a stop, but even during the production of the vowel preceding a stop. Moreover, vowel acoustic duration turned out to be linearly correlated with degree of tongue root advancement, such that longer vowels show greater tongue root advancement.

One of the known differences in supra-glottal articulation between voiced and voiceless stops concerns the position of the tongue root relative to the front-back axis of the oral tract. It has been observed that the tongue root is more advanced in voiced than in voiceless stops (Kent & Moll 1969, Perkell 1969, Westbury 1983). This gesture has been interpreted as a mechanism to ensure voicing can be maintained during the closure of the stop. The realisation of vocal fold vibration (i.e. voicing) requires the air pressure in the supra-glottal cavity to be lower than the air pressure below the glottis. During the production of voiced obstruents, the supra-glottal pressure increases due to the immittance of air in the supra-glottal cavity. Such pressure increase can hinder the ability to sustain voicing during closure, to the point that voicing ceases if the supra-glottal pressure is higher than the sub-glottal pressure (Ohala 2011). One of the possible articulatory solution to counterbalance the increase in pressure during the closure of a voiced stop is to expand the supra-glottal cavity by advancing the root of the tongue.

Tongue root advancement has also been reported as a mechanism for ensuring a short voice onset time (Ahn & Davidson 2016).

An extensive number of studies show that, cross-linguistically, vowels tend to be longer when followed by voiced obstruents than when they are followed by voiceless obstruents (House & Fairbanks 1953, Peterson & Lehiste 1960, Chen 1970, Klatt 1973, Lisker 1974, Farnetani & Kori 1986, Fowler 1992, Hussein 1994, Esposito 2002, Lampp & Reklis 2004, Durvasula & Luo 2012). This phenomenon, know as the voicing effect, has been reported in a variety of languages, including (but not limited to) English, German, Hindi, Russian, Italian, Arabic, and Korean (see Maddieson & Gandour 1976 for a more comprehensive list). A common stance in the literature is that the magnitude of the voicing effect differs depending on the language (although see Laeufer 1992), and that this phenomenon is not a universal tendency, since the duration of vowels is not affected by the voicing of the following obstruents in some languages, like Polish and Czech (Keating 1984). Although several attempts have been made to explain the voicing effect, an account that survives all empirical data is still lacking (Durvasula & Luo 2012, Sóskuthy 2013).

To summarise, tongue root advancement, shorter VOT duration, and longer vowel durations are all correlates of voicing. Moreover, tongue root advancement and shorter VOT show a link. In this paper, I will report the results from an exploratory study which fill the gap in this picture of correlation, by showing that tongue root advancement is also linked to longer vowel durations.

2 Methodology

12 native speakers of Italian (5 females, 7 males) and 6 native speakers of Polish (3 females, 3 males) were recorded in the Phonetics Laboratory at the University of Manchester and in a private location in Italy (see **??**). The Italian speakers of this study are from Northern Italy (three from the north-west and one from north-east). The Polish group was more heterogeneous, with two speakers from the West (Poznań), and two from the East (Warsaw and Przasnysz). Ethical clearance was obtained for this study from the University of Manchester (REF 2016-0099-76). The participants received a monetary compensation.

2.1 Materials

Disyllabic words of the form $C_1V_1C_2V_2$ were used as targets, where C_1 = /p/, V_1 = /a, o, u/, C_2 = /t, d, k, g/, and $V_2 = V_1$ (e.g. *pata*, *pada*, *poto*, etc.), giving a total of 12 target words, used both for Italian and Polish. Most of these words were nonce words in both languages, with a few exceptions (see table). The words were presented using the respective writing conventions (see table). A labial stop was chosen as the first consonant to reduce possible coarticulation with the following vowel.[1] At low power settings and high frame rates, high and mid front vowel have the double disadvantage of being often not clearly visible in the ultrasound image (given their greater distance from the probe) and of producing less displacement of the tongue (essential for the closing gesture identification, see XXX) in the movement from the vowel itself to the following consonant. For this reason, only central/back vowels (low /a/, mid /o/, and high /u/) were included in the target words. The use of back and central vowels (with the exclusion of mid/high front vowels) had the advantage of facilitating the identification of the consonantal gesture of C2. Since the original motive for the exploratory study was to study possible difference in the articulation of closure in voiceless vs. voiced stops, only coronal and velar stops were chosen as target consonants since, of course, the closure of labial consonants cannot be imaged with ultrasonography. The target words were embedded in a frame sentence, *Dico X lentamente* 'I say X slowly' for Italian, and *Mówię X teraz* 'I say X now' for Polish. The similarity of prosodic structure of these sentences ensured better comparability between the two languages.

[1]Although there is a tendency in the articulatory literature to suggest that labial consonants do not affect lingual articulations, Vazquez-Alvarez & Hewlett 2007 report tongue body lowering in the context of labial stops.

2.2 Equipment and procedure

An Articulate Instruments Ltd™ set-up was used for this study (**??**). The ultrasonic data was collected through a TELEMED Echo Blaster 128 unit with a TELEMED C3.5/20/128Z-3 ultrasonic transducer (20mm radius, 2-4 MHz). Stabilisation of the ultrasonic transducer was ensured by using a headset sold by Articulate Instruments Ltd™ (2008). A synchronisation unit (P-Stretch) was plugged into the Echo Blaster unit and used for automatic audio/ultrasound synchronisation. A FocusRight Scarlett Solo pre-amplifier and a Movo LV4-O2 Lavalier microphone were used for audio recording. The acquisition of the ultrasonic and audio signals was achieved with the software Articulate Assistant Advanced (AAA, v2.17.2) running on a Hawlett-Packard ProBook 6750b laptop with Microsoft Windows 7. The synchronisation of the ultrasonic and audio signals was performed by AAA after recording.

The participant's occlusal plane was obtained using a bite plate (Scobbie et al. 2011), and the hard palate was imaged by asking the participant to swallow water (Epstein & Stone 2005). The frame rate of the acquisition of the ultrasonic data ranged between 43 and 68 frames per second. Other settings values varied depending on the frame rate. The ranges of these settings in this study were: scanlines = 88–114, pixel per scanline = 980–988, field of view = 71–93, pixel offset = 109–263, depth (mm) = 75–180. The audio signal was recorded at 22050 MHz (16-bit). The sentences containing the target words were randomised between participants, although the order was kept the same between repetitions within participant due to design constraints of the AAA software. Each participant produced six repetitions of each sentence, which led to a grand total of 1296 sentences.

2.3 Data processing and analysis

The audio data was subject to force alignment using the SPeech Phonetisation Alignment and Syllabification software (SPPAS) (Bigi 2015). The outcome of the automatic alignment was then manually corrected, according to the criteria in Table 1. The onset of the target consonant release (C2 release) was detected automatically in Praat (Boersma & Weenink 2016) by means of the burst-detection algorithm described in Ananthapadmanabha et al. (2014). The durations of the following intervals were extracted from the annotated acoustic landmarks using a scripted procedure in Praat: vowel duration (V1 onset to V1 offset), consonant duration (V1 offset to V2 onset), and closure duration (V1 offset to C2 release).

References

Ahn, Suzy & Lisa Davidson. 2016. Tongue root positioning in English voiced obstruents: Effects of manner and vowel context. *The Journal of the Acoustical Society of America* 140(4). 3221–3221.

Ananthapadmanabha, T. V., A. P. Prathosh & A. G. Ramakrishnan. 2014. Detection of the closure-burst transitions of stops and affricates in continuous speech using the plosion index. *The Journal of the Acoustical Society of America* 135(1). 460–471.

Table 1: List of measurements as extracted from acoustics.

landmark		criteria
vowel onset	(V1 onset)	appearance of higher formants in the spectrogram following the release of /p/ (C1)
vowel offset	(V1 offset)	disappearance of the higher formants in the spectrogram preceding the target consonant (C2)
consonant onset	(C2 onset)	corresponds to V1 offset
closure onset	(C2 closure onset)	corresponds to V1 offset
consonant offset	(C2 offset)	appearance of higher formants of the vowel following C2 (V2); corresponds to V2 onset
consonant release	(C2 release)	automatic detection (Ananthapadmanabha et al. 2014)

Articulate Instruments Ltd™. 2008. Ultrasound stabilisation headset users manual: Revision 1.4. Edinburgh, UK: Articulate Instruments Ltd.

Bigi, Brigitte. 2015. SPPAS - Multi-lingual approaches to the automatic annotation of speech. *The Phonetician* 111–112. 54–69.

Boersma, Paul & David Weenink. 2016. Praat: doing phonetics by computer [Computer program]. Version 6.0.23. http://www.praat.org/.

Chen, Matthew. 1970. Vowel length variation as a function of the voicing of the consonant environment. *Phonetica* 22(3). 129–159.

Durvasula, Karthik & Qian Luo. 2012. Voicing, aspiration, and vowel duration in Hindi. *Proceedings of Meetings on Acoustics* 18. 1–10.

Epstein, Melissa A. & Maureen Stone. 2005. The tongue stops here: Ultrasound imaging of the palate. *The Journal of the Acoustical Society of America* 118(4). 2128–2131.

Esposito, Anna. 2002. On vowel height and consonantal voicing effects: Data from Italian. *Phonetica* 59(4). 197–231.

Farnetani, Edda & Shiro Kori. 1986. Effects of syllable and word structure on segmental durations in spoken Italian. *Speech communication* 5(1). 17–34.

Fowler, Carol A. 1992. Vowel duration and closure duration in voiced and unvoiced stops: There are no contrast effects here. *Journal of Phonetics* 20(1). 143–165.

House, Arthur S. & Grant Fairbanks. 1953. The influence of consonant environment upon the secondary acoustical characteristics of vowels. *The Journal of the Acoustical Society of America* 25(1). 105–113.

Hussein, Lutfi. 1994. *Voicing-dependent vowel duration in Standard Arabic and its acquisition by adult american students*: The Ohio State University dissertation.

Keating, Patricia A. 1984. Universal phonetics and the organization of grammars. *UCLA Working Papers in Phonetics* 59.

Kent, Raymond D. & Kenneth L. Moll. 1969. Vocal-tract characteristics of the stop cognates. *Journal of the Acoustical Society of America* 46(6B). 1549–1555.

Klatt, Dennis H. 1973. Interaction between two factors that influence vowel duration. *The Journal of the Acoustical Society of America* 54(4). 1102–1104.

Laeufer, Christiane. 1992. Patterns of voicing-conditioned vowel duration in French and English. *Journal of Phonetics* 20(4). 411–440.

Lampp, Claire & Heidi Reklis. 2004. Effects of coda voicing and aspiration on Hindi vowels. *The Journal of the Acoustical Society of America* 115(5). 2540–2540.

Lisker, Leigh. 1974. On "explaining" vowel duration variation. In *Proceedings of the Linguistic Society of America*, 225–232.

Maddieson, Ian & Jack Gandour. 1976. Vowel length before aspirated consonants. In *UCLA Working papers in Phonetics*, vol. 31, 46–52.

Ohala, John J. 2011. Accommodation to the aerodynamic voicing constraint and its phonological relevance. In *Proceedings of the 17th International Congress of Phonetic Sciences*, 64–67.

Perkell, Joseph S. 1969. *Physiology of speech production: Results and implication of quantitative cineradiographic study*. Cambridge, MA: MIT Press.

Peterson, Gordon E. & Ilse Lehiste. 1960. Duration of syllable nuclei in english. *The Journal of the Acoustical Society of America* 32(6). 693–703.

Scobbie, James M., Eleanor Lawson, Steve Cowen, Joanne Cleland & Alan A. Wrench. 2011. A common co-ordinate system for mid-sagittal articulatory measurement. In *QMU CASL Working Papers*, 1–4.

Sóskuthy, Márton. 2013. *Phonetic biases and systemic effects in the actuation of sound change*: University of Edinburgh dissertation.

Vazquez-Alvarez, Yolanda & Nigel Hewlett. 2007. The 'trough effect': an ultrasound study. *Phonetica* 64(2-3). 105–121.

Westbury, John R. 1983. Enlargement of the supraglottal cavity and its relation to stop consonant voicing. *The Journal of the Acoustical Society of America* 73(4). 1322–1336.