VOICE ONSET AND OFFSET OSCILLATORY PATTERNS BASED ON PLACE, MANNER AND VOICED/VOICELES CONSONANT ENVIRONMENT

Melda Kunduk[1,2,3], Takeshi Ikuma[1,2], Andrew J McWhorter[1,2], Tobias Schraut[4], Michael Doellinger[4], Robin Samlan[5]

[1]Dept. of Otolaryngology-Head and Neck Surgery, LSU Health Sciences Center, New Orleans, Louisiana, U.S.A.
[2]Voice Center, The Our Lady of The Lake Regional Medical Center, Baton Rouge, Louisiana, U.S.A.
[3]Dept. of Communication Sciences & Disorders, Louisiana State University, Baton Rouge, Louisiana, U.S.A.
[4]Dept. of Otorhinolaryngology Head & Neck Surgery, Medical School Division of Phoniatrics and Pediatric Audiology, University Hospital Erlangen, Erlangen, Germany
[5]Dept. of Speech, Language, & Hearing Sciences, University of Arizona, U.S.A.

Keywords: *Voice; Highspeed Videoendoscopy; Voice onset and offset oscillatory patterns*

Abstract

Objectives / Introduction:

During speech production, the vocal folds (VFs) undergo almost continuous adduction and abduction to generate the voiced and unvoiced segments. Each voiced segment consists of vibratory patterns that include onset and offset and possibly a short duration of vibration. The mechanics and kinematics of the vocal onset and offset of oscillations potentially contain rich information that can be related to vocal health, dysfunction, and patients' specific voice symptoms such as vocal fatigue and handicap (1,2,3). This presentation demonstrates effects of English consonants varied by their place, manner, and voicing on the onset and offset VF oscillation patterns in a vowel-consonant-vowel (VCV) non-word syllable.

Methods:

Data: High speed recordings (HSV) obtained with nasoendoscopy were used to investigate the onset and offset vibratory patterns during the VCV productions, with consonants varying by their place (bilabial, labiodental, alveolar, labiodental, and velar), manner (stop and fricative), and voicing (voiced/voiceless) features (see Table 1). Minimum of six productions of each VCV (/ipi/, /ibi/, /ifi/, /ivi/, /iti/, /idi/, /isi/, /izi/, /iki/, /igi/) were obtained while a female subject produced them with her habitual voice, resulting in 96 HSV recordings total at 4000 frames/second (fps). The rate of the VCV production was regulated by a metronome to ensure uniformity across recordings (0.92 seconds per production). HSV images and acoustic data were recorded simultaneously.

Procedures: HSV recordings were temporally segmented using the acoustic data as a guide, and each VCV production was individually saved to a video file to play back at 20 fps (i.e., slowed down by 200). Each video was perceptually evaluated around its consonant for the presence or absence of three outcome measures: Arytenoid Abduction (separation), Vocal Fold Vibration Cessation, and Vocal Fold Contact Loss.

Analysis: The data were analyzed in 3 two-factor groups while fixing either manner or place: voicing and place effects on stops (Group 1, N:36); voicing and place effects on fricatives (Group 2; N:28); and voicing and manner effects on alveolar consonants (Group 3, N:28). This arrangement avoids a missing-data problem due to English lacking some consonants (i.e., /ɸ/, /β/, /p̺/, /b̺/, /x/, and /ɣ/). The numbers of samples in VCV groups were matched to the minimum available. Those with more samples were randomly sampled for each group.

Table 1: Investigated Vowel-Consonant-Vowel Patterns and Analysis Groupings

Manner	Voicing	Place: Bilabial	Place: Labiodental	Place: Alveolar	Place: Velar	
Stop	Voiceless	/ipi/		/iti/	/iki/	⇦ *Group 1*
	Voiced	/ibi/		/idi/	/igi/	
Fricative	Voiceless		/ifi/	/isi/		⇦ *Group 2*
	Voiced		/ivi/	/izi/		
				⇧ *Group 3*		

Results:

Group 1: Voicing and place effects for stops: All voiceless stops (/ipi/, /iti/, /iki/; N:18) (Table 2) presented with arytenoid abduction, vocal fold vibration cessation, and vocal fold contact loss across alveolar, bilabial, and velar places of articulation before the stop consonant in VCV production. None of the voiced stops (/ibi/, /idi/, /igi/; N:18) (Table 2) consonants demonstrated arytenoid abduction before stopping consonant production. Vocal fold vibration cessation was present only with voiced bilabial and alveolar stops (/ibi/, /idi/); both were present only once in six productions (17%). Vocal fold contact loss happened in 39% of all voiced stops while 83% of voiced bilabial and 33% of voiced alveolar stop productions presented this feature (Table 2).

Group 2: Voicing and place effects for fricatives: All voiceless fricative production (/ifi/, /isi/; N:14) (Table 3) presented arytenoid abduction, vocal fold vibration cessation, and vocal fold contact loss for labiodental and alveolar places of articulation before the fricative production in VCV production. None of the voiced fricatives (/ivi/, /izi/; N:14) demonstrated arytenoid abduction before the fricative in VCV production. Vocal fold vibration cessation was present only in voiced alveolar fricatives (/izi/; N:7) indicating place of articulation effect during VCV production. None of the voiced labiodental fricatives demonstrated this gesture in VCV production. Vocal fold contact loss occurred in all alveolar voiced fricatives (/izi/; N:7) and only 14% voiced labiodental fricatives (/ivi/; N:7) indicating a place of articulation effect during VCV production (Table 3).

Group 3: Voicing and manner effects on alveolar place of articulation: While all voiceless alveolar stops and fricatives (/iti/, /isi/; N:14) demonstrated vocal fold contact loss before the consonant production, only 43% of all voiced alveolar stops (/idi/) and 100% voiced alveolar fricatives (/izi/) did so in VCV production (Table 4).

Table 2: Group 1 / *Voicing and Place Effects* on Stop Consonants

Voicing	Place	Arytenoid Abduction	VF Vibration Cessation	Vocal Fold Contact Loss
Voiceless	(all)	18/18 (100%)	18/18 (100%)	18/18 (100%)
Voiced	Bilabial	0/6 (0%)	1/6 (17%)	5/6 (83%)
	Alveolar	0/6 (0%)	1/6 (17%)	2/6 (33%)
	Velar	0/6 (0%)	0/6 (0%)	0/6 (0%)

Table 3: Group 2 / *Voicing and Place Effects* on Fricative Consonants

Voicing	Place	Arytenoid Abduction	VF Vibration Cessation	Vocal Fold Contact Loss
Voiceless	(all)	14/14 (100%)	14/14 (100%)	14/14 (100%)
Voiced	Labiodental	0/7 (0%)	0/7 (0%)	1/7 (14%)
	Alveolar	0/7 (0%)	7/7 (100%)	7/7 (100%)

Table 4: Group 3 / *Voicing and Manner* Effects on Alveolar Consonants

Voicing	Manner	Arytenoid Abduction	VF Vibration Cessation	Vocal Fold Contact Loss
Voiceless	(all)	14/14 (100%)	14/14 (100%)	14/14 (100%)
Voiced	Stop	0/7 (0%)	1/7 (14%)	3/7 (43%)
	Fricative	0/7 (0%)	7/7 (100%)	7/7 (100%)

Conclusions:

Preliminary results strongly suggest main effects of voicing, place, and manner and interaction effects between voicing and place and between voicing and manner on the vocal fold vibratory patterns of the VCV productions. Arytenoid abduction (separation) with cessation of vibration and loss of vocal fold contact always occurs during VCV production with all of the voiceless consonants assessed (/ipi/, /iti/, /iki/, /ifi/, /isi/). As the arytenoids abduct, VFs have to lose contact, and there is too much air leakage to sustain the vibration. None of the voiced consonants (/ibi/, /idi/, /igi/, /ivi/, /izi/) exhibited arytenoid separation. The vocal fold vibration cessation before the voiced consonant production was observed mostly for alveolar fricative (/izi/) in VCV position. The vocal fold contact loss was observed in some of the voiced alveolar stops but present in all voiced alveolar fricatives. Behaviors of the vocal fold (contact/cessation) with the voiced consonants is dictated by the aerodynamic coupling of the source and vocal tract. Interestingly, the stops and fricatives draw the polar opposite effect of the placement: stops further from source cause more contact loss while fricatives further from source cause less contact loss and cessation. At the alveolus, the voiced fricative is more likely to cause disruption to the vibration than voiced stop. Indications for future studies and clinical implications of the results will be discussed.

References

1. M. Kunduk, T. Ikuma, D. C. Blouin, and A. J. McWhorter, “Effects of volume, pitch, and phonation type on oscillation initiation and termination phases investigated with high-speed videoendoscopy,” J. Voice, vol. 31, no. 3, pp. 313–322, 2017, doi: 10.1016/j.jvoice.2016.08.016.
2. R. F. Orlikoff, D. D. Deliyski, R. J. Baken, and B. C. Watson, “Validation of a glottographic measure of vocal attack,” J. Voice, vol. 23, no. 2, pp. 164–168, 2009, doi: 10.1016/j.jvoice.2007.08.004.
3. C. E. Stepp, “Relative fundamental frequency during vocal onset and offset in older speakers with and without Parkinson’s disease,” J. Acoust. Soc. Am., vol. 133, no. 3, pp. 1637–1643, Mar. 2013, doi: 10.1121/1.477620