Temporal stability and compensatory adjustments: Data on the effect of voicing on vowel duration in Italian and Polish

The duration of vowels has been shown to be affected by the voicing of the following consonant in several languages (House & Fairbanks, 1953; Lisker, 1974; Laeufer, 1992). There is a strong tendency for vowels followed by voiced stops to be long, and for vowels followed by voiceless stops to have shorter durations. The source of this so-called 'voicing effect' is still subject to disagreement. The compensatory temporal adjustment account (Slis & Cohen, 1969; Lehiste, 1970) argues that, given a temporally stable speech interval like the syllable or the word, an increase in consonant duration will correspond to a decrease in vowel duration, and vice versa. Voiceless stops tend to have longer closure durations than voiced stops, so that a longer closure duration would be preceded by a shorter vowel (Lisker, 1957). Arguments against the compensatory account have been put forward in light of a failure to find a stable interval and considering empirical data on other durational effects (Chen, 1970; Maddieson & Gandour, 1976). In this paper, I discuss the results from an exploratory study of Italian and Polish which are compatible with a compensatory account, and I argue that the temporally stable interval in stressed syllables involving stops in these languages is the interval between two consecutive stop releases (the 'release-to-release' interval).

Several aspects of the experiment design were constrained by the use of ultrasound tongue imaging, such as the choice of vowels and the number of repetitions. The ultrasound data will not be discussed here since it is not of immediate relevance to the argument put forward. Audio recordings were obtained from 11 speakers of Italian (5F) and 6 speakers of Polish (3F). Nonce words were used as target words, constructed according to the structure $C_1\acute{V}_1C_2V_2$, 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.). The resulting 12 words were embedded in a frame sentence, *Dico* ___ *lentamente* 'I say ___ slowly' in Italian, and *Mówię* ___ *teraz* 'I say ___ now'. Each participant repeated the 12 stimuli six times. The duration of the following intervals were extracted from acoustic landmarks: V_1, C_2 closure, C_1 release to C_2 release.

According to linear mixed-effects models, vowels are 16 ms (se = 4.4) longer when followed by a voiced stop both in Italian and Polish (the interaction between voicing and language is not significant, although this might be related to low statistical power). Stop closure is 17 ms (se = 4) shorter if the stop is voiced. Closure and vowel duration are weakly negatively correlated (β = -0.19 ms, se = 0.06 ms). A Bayes factor analysis suggests that a null effect of C_2 voicing on the release-to-release duration is 23 times more likely than the presence of an effect. This is interpreted as positive evidence (Raftery, 1999) that the release-to-release interval duration is not affected by C_2 voicing.

The results of this study are in agreement with a compensatory temporal account of the voicing effect (although other articulatory and perceptual factors cannot be ruled out and they might still contribute to the overall effect). Since the interval between two consecutive releases is not affected by the voicing of the stop following the vowel, under a compensatory account the timing of the closure onset decides the respective durations of the vowel and the closure. Since the closure of voiced stops is initiated later within the release-to-release interval, the preceding vowel is longer when compared to a vowel followed by a voiceless stop, the closure of which starts earlier. These patterns are also compatible with Fowler's (1983) approach to the cyclic production of vocalic gestures and with the effects on vowel duration of differing numbers of following consonants (Farnetani & Kori, 1986). Although to be confirmed with data from other phonological contexts and languages, this temporal pattern has implications for other laryngeal effects (like aspiration and ejection, Maddieson & Gandour 1976; Beguš 2017) and for gestural phasing (Tilsen, 2013, 2016). These implications will be discussed along with future directions.

References

Beguš, Gašper. 2017. Effects of ejective stops on preceding vowel duration. *The Journal of the Acoustical Society of America* 142(4). 2168–2184. doi:10.1121/1.5007728.

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

Farnetani, Edda & Shiro Kori. 1986. Effects of syllable and word structure on segmental durations in spoken Italian. *Speech communication* 5(1). 17–34. doi:10.1016/0167-6393(86)90027-0.

Fowler, Carol A. 1983. Converging sources of evidence on spoken and perceived rhythms of speech: Cyclic production of vowels in monosyllabic stress feet. *Journal of Experimental Psychology: General* 112(3). 386. doi:10.1037/0096-3445.112.3.386.

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. doi:10.1121/1.1906982.

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

Lehiste, Ilse. 1970. Temporal organization of higher-level linguistic units. *The Journal of the Acoustical Society of America* 48(1A). 111–111. doi:10.1121/1.1974906.

Lisker, Leigh. 1957. Closure duration and the intervocalic voiced-voiceless distinction in English. *Language* 33(1). 42–49. doi:10.2307/410949.

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.

Raftery, Adrian E. 1999. Bayes factors and BIC: Comment on "A critique of the Bayesian information criterion for model selection". *Sociological Methods & Research* 27(3). 411–427. doi:10.1177/0049124199027003005.

Slis, Iman H. & Antonie Cohen. 1969. On the complex regulating the voiced-voiceless distinction II. *Language and speech* 12(3). 137–155. doi:10.1177/002383096901200301.

Tilsen, Sam. 2013. A dynamical model of hierarchical selection and coordination in speech planning. *PLoS ONE* 8(4). e62800. doi:10.1371/journal.pone.0062800.

Tilsen, Sam. 2016. Selection and coordination: The articulatory basis for the emergence of phonological structure. *Journal of Phonetics* 55. 53–77. doi:10.1016/j.wocn.2015.11.005.