

Exploring the connection between schwa dropping and vowel lengthening in Albanian

Compensatory lengthening (CL) is a phonological process where a target segment is lengthened after the loss of another, the trigger (see e.g. Gess 2011; Kavitskaya 2002). The present study examines an emergent case of CVCV > CV:C type of CL in Albanian, a language of the Indo-European family which comprises two main dialects, Gheg and Tosk. A phonological difference between these dialects is that Gheg has vowel quantity contrasts, but not Tosk (except for a few speech communities not covered here; Demiraj 1996; Gjinari et al. 2007; Totoni 1964). Contrastive length is found in Gheg, among other contexts, in definite versus indefinite singular nouns, which have short and long stem vowels respectively, while indefinites are marked in Tosk by a final unstressed schwa, as in example (1) (Çeliku 1971, 2020). Historically, these long vowels have been argued to have developed in Gheg because of CL triggered by schwa dropping (e.g. Çabej 1957; Çeliku 1971).

(1)	<i>buka</i> ‘the bread’ (definite)	<i>një bukë</i> ‘a bread’ (indefinite)
Gheg	/buka/	/bu:k/
Tosk	/buka/	/bukə/

Parting from classic descriptions of Tosk, Çeliku (1971) noted cases of schwa dropping in indefinite nouns by Tosk speakers, with variation within and across individuals suggesting change in progress. The main aim of this study is to find out whether CL is emerging in indefinite nouns in Tosk following schwa dropping, similarly to the development observed in Gheg. However, the debated claim that contrastive length must already exist in a language in order for CL to occur could perhaps restrict its emergence in Tosk (de Chene & Anderson 1979; Kavitskaya 2017). Furthermore, Çeliku (1971) noted that contrary to the fairly common CL process, Tosk nouns *without* schwa had a *short* stem vowel, while nouns *with* schwa presented *longer* stem vowels, an impressionistic claim which this study aims at revisiting empirically.

The participants were 16 speakers of Tosk (5F, 11M, 20-75 y.o., \bar{x} = 41 y.o.) and 17 speakers of Gheg (9F, 8M, 38-79 y.o., \bar{x} = 55 y.o.) recorded during field trips in Albania in 2020. They took part in a reading task where two repetitions of a target word were inserted in a carrier sentence, e.g. *buka thoni buka* ‘the bread I say the bread’. Each sentence was presented twice in a random order. The recordings were forced-aligned using WebMAUS (Kisler et al. 2017) then converted into an EMU speech database (Winkelmann et al. 2017) for manual correction of the segment boundaries. The material selected for this paper consists of 9 pairs of definite/indefinite singular nouns such as *buka/bukë* ‘the/a bread’, *kali/kalë* ‘the/a horse’, etc. Two separate linear mixed-effect regression models were used to test how the log-transformed duration of the stem vowels was affected by: a) the presence or absence of a word-final schwa in indefinite nouns uttered by Tosk speakers; b) definiteness and dialect.

Descriptive statistics for Tosk indicate that 78% of the indefinite nouns lacked schwa and that no speaker categorically produced it. The results of analysis a) (see Fig.1) show no significant difference between indefinite nouns with and without schwa. The results of analysis b) (see Fig.2) first indicate that Gheg speakers produced vowels with a significantly longer duration than Tosk speakers irrespective of definiteness. Second, a main effect of definiteness was also found, with stem vowels in indefinite nouns being significantly longer than in definite nouns across dialects. Taken together, these results support to a certain extent the hypothesis of emergent CL despite the lack of quantity contrasts in Tosk: we observe similar patterns in Tosk and Gheg where vowels in indefinite nouns have a longer duration than in definite nouns, a tendency for schwa dropping in Tosk, as well as no empirical evidence of the pattern opposite to CL described by Çeliku (1971). The absence of a clear effect of schwa dropping on vowel length does not point to a phonetic process of physical duration preservation (Gess 2011).

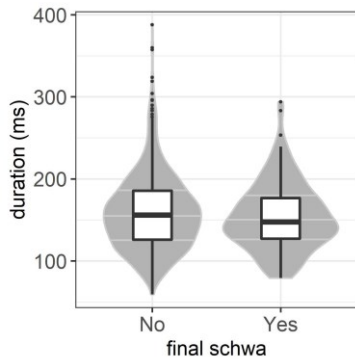


Fig.1: Violin and box plots of duration of stem vowels in indefinite nouns produced by Tosk speakers, depending on the presence (Yes) or absence (No) of a final schwa

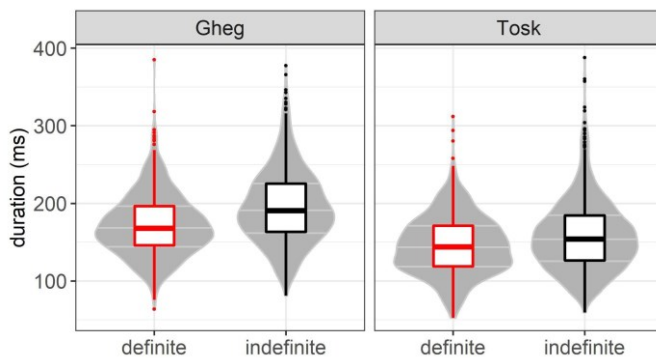


Fig.2: Violin and box plots of duration of stem vowels in definite (red) versus indefinite (black) nouns produced by Gheg (left) and Tosk (right) speakers

References

- Çabej, E. (1957). Kuantiteti i zanoreve të theksuara të shqipes. *Buletin i Universitetit Shtetëror Të Tiranës. Seria Shkenca Shoqërore*, 2.
- Çeliku, M. (1971). Kuantiteti i zanoreve të theksuara në të folmet e shqipes. *Studime Filologjike*, 8(4), 65–100.
- Çeliku, M. (2020). *Gegërishtja jugperëndimore*. Akademia e Studimeve Albanologjike.
- de Chene, B. E., & Anderson, S. R. (1979). Compensatory lengthening. *Language*, 55(3), 505–535.
- Demiraj, S. (1996). *Fonologjia historike e gjuhës shqipe*. Akademia e Shkencave e Republikës së Shqipërisë.
- Gess, R. (2011). Compensatory lengthening. In M. van Oostendorp, C. J. Ewen, E. Hume, & K. Rice (Eds.), *The Blackwell Companion to Phonology: Vol. III* (pp. 1536–1557). Wiley Blackwell.
- Gjinari, J., Beci, B., Shkurtaj, G., Gosturani, X., & Dodi, A. (2007). *Atlasi dialektologjik i gjuhës shqipe*. Università degli Studi di Napoli.
- Kavitskaya, D. (2002). *Compensatory Lengthening: Phonetics, Phonology, Diachrony*. Routledge.
- Kavitskaya, D. (2017). Compensatory lengthening and structure preservation revisited yet again. In C. Bower, L. Horn, & R. Zanuttini (Eds.), *On Looking into Words (and Beyond)* (pp. 41–58). Language Science Press.
- Kisler, T., Reichel, U. D., & Schiel, F. (2017). Multilingual processing of speech via web services. *Computer Speech & Language*, 45, 326–347.
- Totoni, M. (1964). E folmja e Bregdetit të Poshtëm. *Studime Filologjike*, 1(1), 129–158.
- Winkelmann, R., Harrington, J., & Jänsch, K. (2017). EMU-SDMS: Advanced speech database management and analysis in R. *Computer Speech & Language*, 45, 392–410.