

## A longitudinal study of contrastive length in Albanian-speaking children

Enkeleida Kapia<sup>1,2</sup>, Josiane Riverin-Coutlée<sup>1</sup>, Conceição Cunha<sup>1</sup> & Jonathan Harrington<sup>1</sup>

<sup>1</sup>Institute of Phonetics and Speech Processing, Ludwig-Maximilians-Universität München, Germany

<sup>2</sup>Academy of Albanological Sciences, Albania

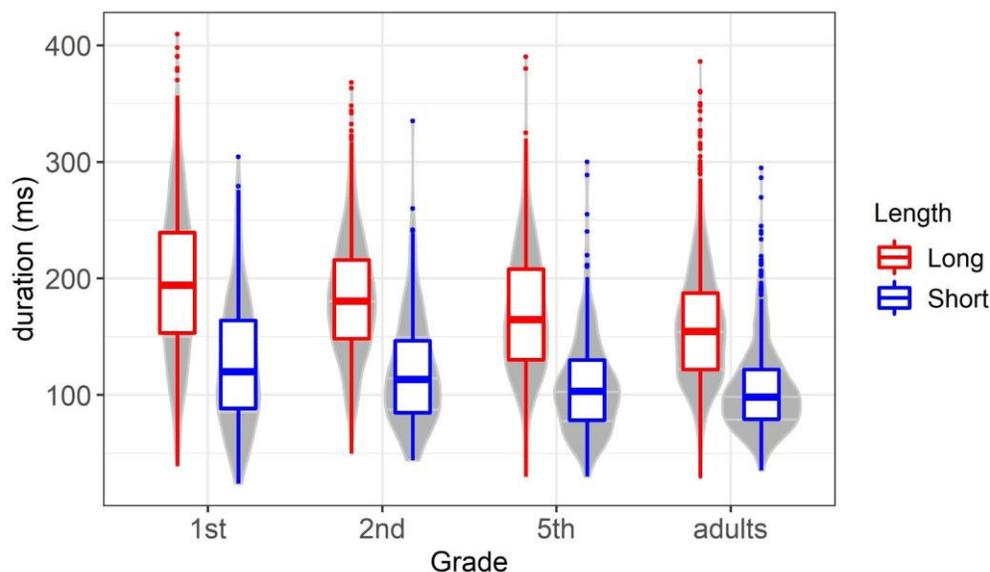
This longitudinal study examines whether contrastive length is lost or reduced under the influence of standard Albanian in Gheg-speaking children during primary school. Albanian is a language of the Indo-European family spoken by approximately 7 million people living mostly in Albania, Kosovo and North Macedonia [8]. Gheg, one of its main dialects, features contrastive vowel length, while standard Albanian does not [2]. Over the past 50 years, there have been impressionistic reports of contrastive length disappearing from Gheg, especially in the greater Tirana area, under the influence of standard Albanian [3]. In a recent study [7], we did not find empirical evidence of the loss of contrastive length in Tirana Gheg: neither in adult participants, nor in a cohort of children recorded in 1<sup>st</sup> and 2<sup>nd</sup> grades (6 and 7 years old). It can be argued, however, that these young children were still replicating the length patterns of their family and had not been influenced by standard Albanian yet [4], while increased exposure to the standard as they progress through schooling may affect their productions later on. Here, we examine data from 10 children of the same cohort when they are in 5<sup>th</sup> grade (10 years old). Our question is whether an influence of standard Albanian can be observed on these children's productions of length contrasts in 5<sup>th</sup> grade when compared to 1<sup>st</sup> and 2<sup>nd</sup> grades.

The 10 participants were recorded in primary schools of the greater Tirana area. They were screened for dialect background and language impairment. They participated in a picture-naming task in which 6 and 19 images depicted words with short and long stressed vowels respectively. The speech signal was forced-aligned using WebMAUS [5] then structured into a database using EMU-SDMS [10]. Segment boundaries were manually corrected in order to measure vowel duration, which was then log-transformed and used as response variable in a linear mixed-effect regression model [1] with *Length* (2 levels) and *School Year* (3 levels) set as fixed effects.

Figure 1 shows how the duration of short and long vowels evolves from 1<sup>st</sup> to 5<sup>th</sup> grades. A main effect of *Length* was found, with significantly greater durations in long than short vowels. No effect of the *School Year* was found, i.e. the small reduction of durations over time that is visible in Figure 1 is not significant. No interaction between *Length* and *School Year* was found either, meaning that the relationship between short and long vowels remained stable over time. These results suggest no attrition of vowel length contrasts in Gheg, unlike other features that differ from standard Albanian and other varieties [7]. This could be due to this feature's involvement in marking morphology (e.g. definite and indefinite singular nouns have short and long vowels respectively), making it particularly entrenched in speakers' grammars and resistant to change. Furthermore, while some speech features are judged impressionistically to reach adult-like proficiency by the age of 10 [9], looking at fine acoustic features like duration shows that speech motor control is still being gradually refined [6].

## References

- [1] Bates, Douglas, Martin Maechler, Ben Bolker & Steve Walker. 2015. Fitting linear mixed-effects models using lme4. *Journal of Statistical Software* 67(1). 1–48.
- [2] Beci, Bahri. 2004. *Fonetika e gjuhës shqipe*. Tiranë: EDFA.
- [3] Çeliku, Mehmet. 2020. *Gegërishtja jugperëndimore*. Tiranë: Akademia e Studimeve Albanologjike.
- [4] Foulkes, Paul & Marilyn Vihman. 2015. First language acquisition and phonological change. In Patrick Honeybone & Joseph Salmons (eds.), *The Oxford Handbook of Historical Phonology*, 289–312. Oxford: Oxford University Press.
- [5] Kisler, Thomas, Uwe D. Reichel & Florian Schiel. 2017. Multilingual processing of speech via web services. *Computer Speech & Language* 45. 326–347.
- [6] Lee, Sungbok, Alexandros Potamianos & Shrikanth Narayanan. 1999. Acoustics of children's speech: Developmental changes of temporal and spectral parameters. *Journal of the Acoustical Society of America* 105(3). 1455–1468.
- [7] Riverin-Coutlée, Josiane, Conceição Cunha, Enkeleida Kapia & Jonathan Harrington. 2021. Dialect features in heterogeneous and homogeneous Gheg speaking communities. In *Proceedings of Interspeech 2021*, 1449–1453. Brno, Czech Republic.
- [8] Rusakov, Alexander. 2017. Albanian. In Mate Kapović (ed.), *The Indo-European Languages*, 552–608. 2nd edn. London: Routledge.
- [9] Schölderle, Theresa, Elisabeth Haas & Wolfram Ziegler. 2020. Age norms for auditory-perceptual neurophonetic parameters: A prerequisite for the assessment of childhood dysarthria. *Journal of Speech, Language, and Hearing Research* 63(4). 1071–1082.
- [10] Winkelmann, Raphael, Jonathan Harrington & Klaus Jänsch. 2017. EMU-SDMS: Advanced speech database management and analysis in R. *Computer Speech & Language* 45. 392–410.



**Figure 1.** Duration of vowels produced by a cohort of 10 children in 1<sup>st</sup>, 2<sup>nd</sup> and 5<sup>th</sup> grades, as well as a control group of 28 Gheg-speaking adults for comparison