

Acquisition of Morphologically and Phonologically Conditioned Vowel Length in Albanian

Enkeleida Kapia

Josiane Riverin-Coutlée

Conceição Cunha

Jonathan Harrington





- Albanian and Gheg
- Vowel lengthening rules in Gheg
- Acquisition of morpho-phonological rules
- Our study (part of a bigger project)
- Results from children and adults
- Discussion in relation to field
- Limitations and future directions

Albanian

- Language spoken in SE Europe
- Considered morpho-phonologically “complex”
(e.g. Hubbard, 1985; Buchholz, 1987)
 - *dashi, e dashit, dashit, dashin, dashit*
(5 declensions of *ram* – def., sing.)
 - *deshtë, e deshëve, deshëve, deshtë, deshëve*
(5 declensions of *rams* – def., plu.)
- Tosk and Gheg dialects
 - **Gheg** (spoken in northern & central Albania & Kosovo)



Elsie & Gross (2009)



Vowel lengthening in Gheg

- Gheg has two types of vowel lengthening rules (e.g. Beci, 1995; Gjinari et al, 2007; Lloshi, 2009; Shkurtaj, 2004):
 1. Morphologically sensitive lengthening (hereafter, **Type 1**)
 2. Phonologically sensitive lengthening (**Type 2**)

Morphologically-conditioned lengthening in Gheg

- **Type 1** characterizes indefinite nouns, as opposed to definite nouns (e.g. /veza/ “the egg” vs. /ve:z/ “an egg”)



/veza/ e bardhë
the white egg



një /ve:z/ e thyer
a broken egg

- Type 1 is the **most frequent** lengthening process (e.g. Shkurtaç, 2004; Çeliku, 1971; Murati, 1989).

Phonologically-conditioned lengthening in Gheg

- **Type 2** happens when vowels occur:

(a) before sonorants



/ka:l/ (horse)



/kat/ (storey, as in one-storey house)

(b) in open final syllables



/mi:/ (mouse)



/miz/ (fly)



Tendencies from previous acquisition studies:

- 1. Statistical frequency** is one of the main predictors of ease and speed of acquisition of morpho-phonological patterns
(e.g. Stemberger, 1993; Demuth, 2007; Ambridge et al., 2015; Tessier, 2016; etc.)
→ Type 1 is more frequent
- 2. Phonological rules** acquired earlier than morphological ones, even if they are less frequent
(e.g. Eimas, 1971; MacWhinney, 1975, 1978; Łukaszewicz, 2006; etc.)
→ Type 2 is a phonological rule



- Discover the developmental pattern of the two different vowel length rules (Type 1 & Type 2) in Gheg speaking children
- Determine whether children benefit more from a frequently occurring morphological rule (Type 1) or a less frequently occurring phonological rule (Type 2) during acquisition
 - Type 1 acquired earlier according to the statistical learning thesis
 - Type 2 acquired earlier according to the phonological rule thesis



Speakers

- Native speakers of Gheg
- 22 adults (29-74 years old, mean = 44.2, 20 women)
 - living in greater Tirana area (village and city)
 - they were the children's parents, grandparents or teachers
- 37 children (6-7 years old, 20 girls)
 - living in greater Tirana area (village and city)
 - earliest age group likely to show phonetic proficiency in communicating morpho-phonologically complex factors in languages like Albanian (e.g. MacWhinney, 1978 for Hungarian; Tomas et al., 2017 for Russian)



Task

- Speakers recorded in primary schools in Albania
(Speech Recorder, Draxler & Jänsch, 2004)
- Picture-naming task designed for 6-7 year olds (1st grade)
- Images of relevant cultural objects, presented 4 times each
- Each image corresponded to a monosyllabic word:
 - 7 words of Type 1
 - 6 words of Type 2
 - 5 control words

Examples of pictures

Type 1



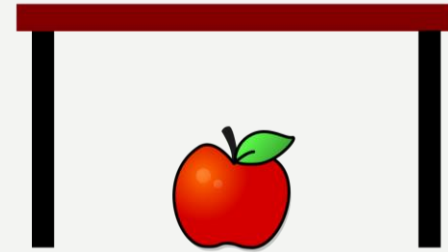
/bu:k/ (*bread*)

Type 2



/mi:/ (*mouse*)

Control

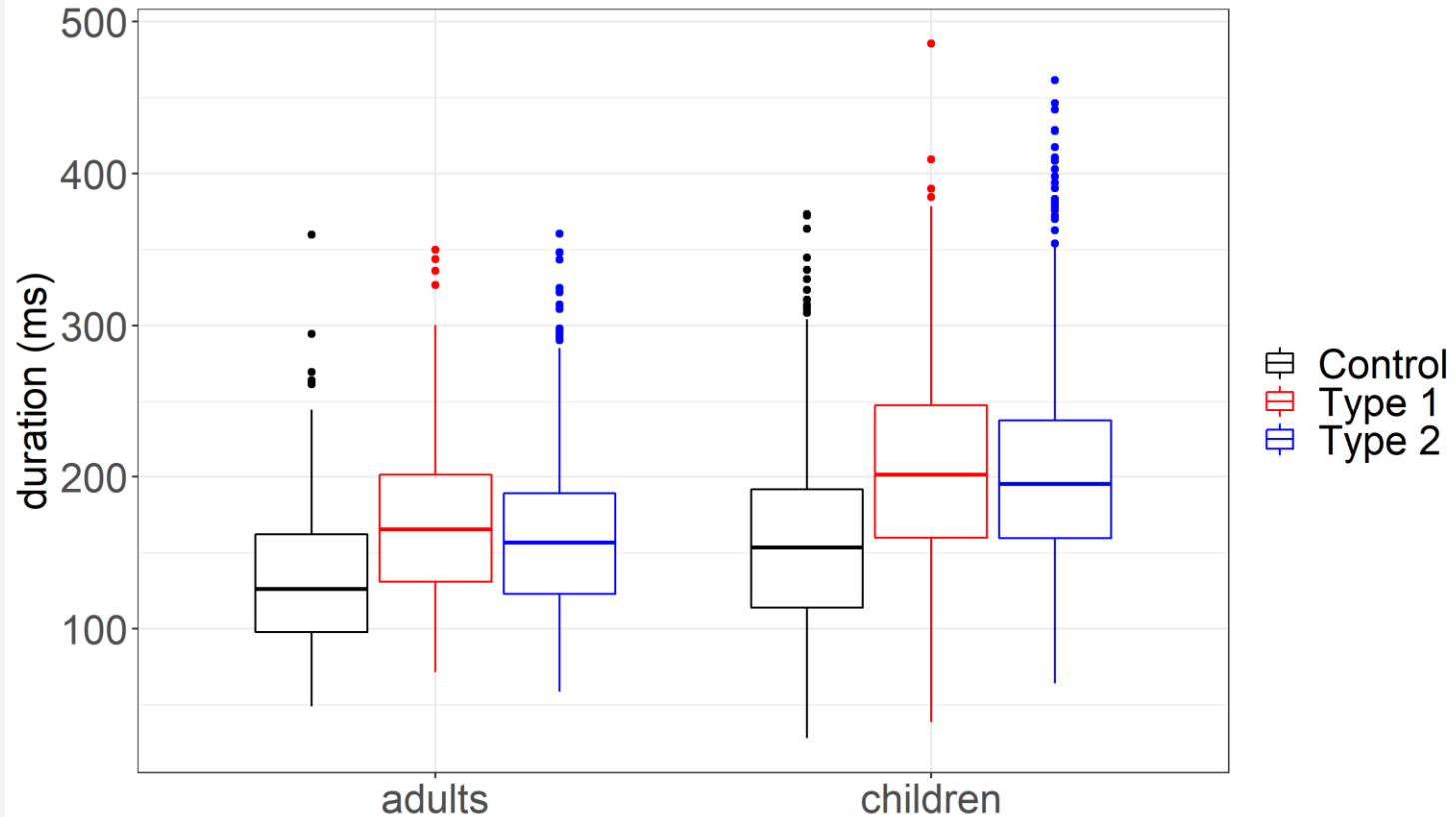


/posht/ (*under*)



Data Processing and Analysis

- Speech signal forced-aligned using WebMAUS (Schiel, 1999; Kisler et al., 2017)
- Database handled in EMU-SDMS (Winkelmann et al., 2017)
 - including hand-correction of segment boundaries
- Statistical analyses with *lme4* and *lmerTest* packages in R (Bates et al., 2015; Kuznetsova et al., 2017; R Core Team, 2020)



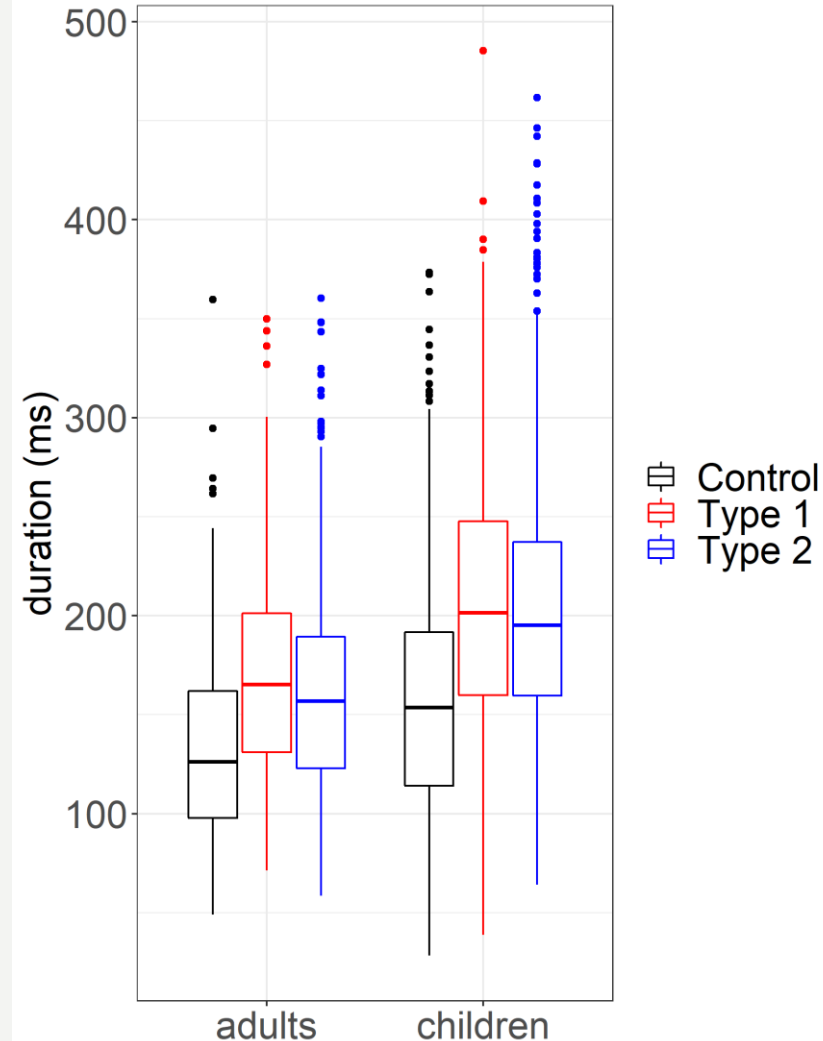
Duration of vowels in Control, Type 1 and Type 2 words in adult and child speakers

Model initially tested: $\text{duration} \sim (\text{word_type} * \text{age_group}) + (\text{age_group} | \text{word}) + (\text{word_type} | \text{speaker})$

Best model found: $\text{duration} \sim \text{word_type} + \text{age_group} + (\text{age_group} | \text{word}) + (\text{word_type} | \text{speaker})$

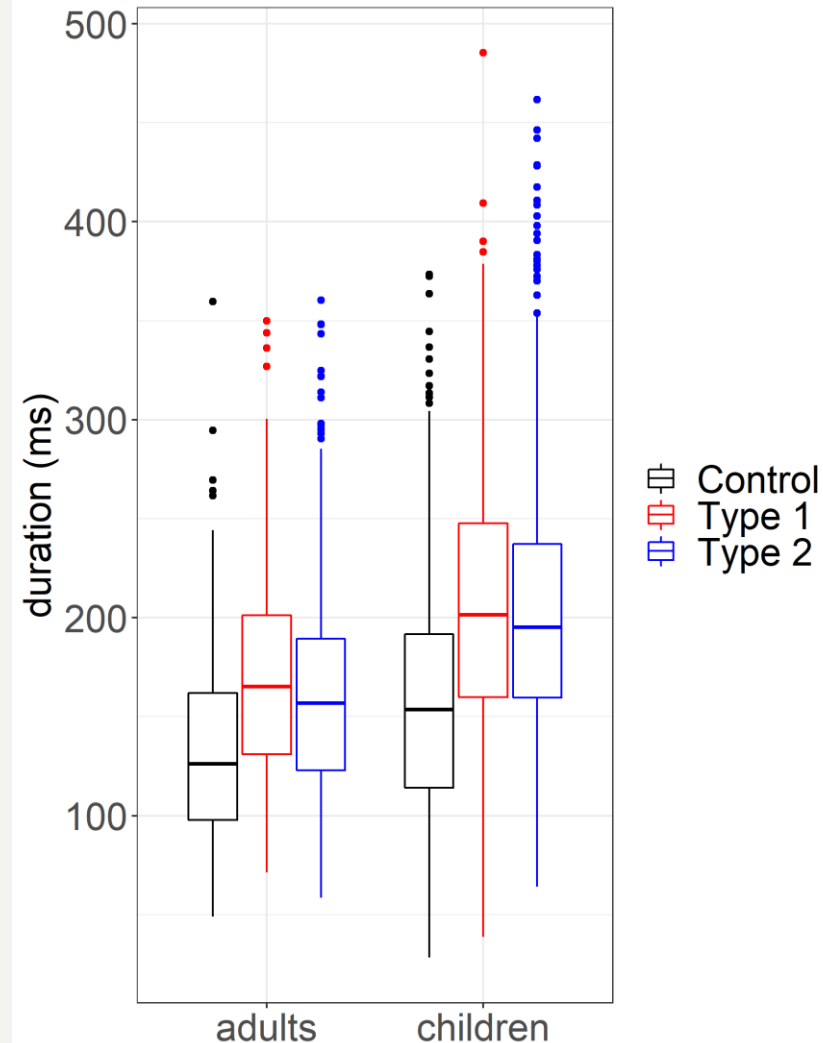


1. Children (right) have longer durations than adults (left)
($F[1, 68.71]=10.25, p<0.01$)
→ Likely due to slower articulation
2. Type 1 (red) and Type 2 (blue) are longer than control (black)
(Type 1: $t[15.9]=3.22, p<0.05$;
Type 2: $t[16.1]=2.88, p<0.05$)
→ Evidence for contrastive vowel lengthening





3. Type 1 (red) and Type 2 (blue) are undifferentiated
($t[15.6]=0.24, p=0.96$)
→ No statistical difference
4. No interaction between word type and age group
($F[2, 23.43]=0.82, p=0.44$)
→ Relationship between word types is the same for children and adults





Taken together, these results show that:

1. Children have mastered both types of lengthening by age 6-7
 - They produce the same length patterns as adults
 - Type 1 and Type 2 are longer than Control words
2. Less frequent phonological and more frequent morphological factors are learned with equal proficiency
 - Type 1 (frequent) and Type 2 (phonological) rules are not different from one another



Limitations and future work

1. Age limitations

- Despite recent findings for Russian (Tomas et al., 2017), perhaps our children are too old for a **developmental** study?
 - Younger age group(s) needed
 - Could be specific to Gheg that length is particularly prominent

2. Sociolinguistic factors

- Age range in adults (parents/teachers vs. grandparents)
- Potential differences between city (Tirana) and village

3. More words and identical phonetic environment needed for a full-fledged study



- Ambridge, B., Kidd, E., Rowland, C.F. & Theakston, A.L. (2015). The ubiquity of frequency effects in first language acquisition. *Journal of Child Language*, 42, 239-273.
- Bates, D., Maechler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, 67(1), 1-48.
- Beci, B. (1995). *Të folmet veriperëndimore të shqipes dhe sistemi fonetik i të folmes së Shkodrës*. Akademia e Shkencave. Tiranë.
- Buchholz, O. & Fiedler, W. (1987). *Albanische Grammatik*. Leipzig: VEB Verlag Enzyklopädie.
- Çeliku, M. (1971). Kuantiteti i zanoreve të theksuara në të folmet e shqipes. *Studime Filologjike*, 4, 65-100. Tiranë.
- Demuth, K. (2007). The role of frequency in language acquisition. In Gülzow, I. & Gagarina, N. (Eds.), *Frequency effects in language acquisition*, (pp. 383-388). Berlin. Mouton De Gruyter.
- Draxler, C. & Jänsch, K. (2004). SpeechRecorder – A universal platform independent multi-channel audio recording software. *Proceedings of the 4th International Conference on Language Resources and Evaluation*, Lisbon, 559-562.
- Eimas, P.D., Siqueland, E.R., Jusczyk, P. & Vigorito, J. (1971). Speech perception in infants. *Science*, 71, 303-306.
- Elsie, R., & Gross, J. (2009). *Albanian dialects*. Robert Elsie. Writer, Translator, Interpreter, Specialist in Albanian Studies (1988-2019). <http://www.elsie.de/>
- Gjinari, J., Beci, B., Shkurtaç, Gj., Gosturani, Xh. & Dodi, A. (2007). *Atlasi Dialektologjik i Gjuhës Shqipe*. Universita degli studi di Napoli L'Orientale, Dipartimento di studi dell'Europa orientale, Napoli.
- Hubbard, P. (1985). *The Syntax of the Albanian Verb Complex*. New York–London: Garland Publishing.
- Kisler, T., Reichel, U. D. & Schiel, F. (2017). Multilingual processing of speech via web services. *Computer Speech & Language*, 45, 326-347.
- Kuznetsova, A., Brockhoff, P. B. & Christensen, R. H. B. (2017). lmerTest package: Tests in linear mixed effects models. *Journal of Statistical Software*, 82(13), 1-26.
- Lloshi, Xh. (1999). Albanian. In U. Hindrichs (Ed.), *Handbuch der Südosteuropa-Linguistik*, (pp. 277-299). Harrassowitz Verlag.
- Łukaszewicz, B. (2006). Extrasyllabicity, transparency and prosodic constituency in the acquisition of Polish. *Lingua*, 116, 1-30.
- MacWhinney, B. (1975). Rules, rote, and analogy in morphological formations by Hungarian children. *Journal of Child Language*, 2, 65-77.
- MacWhinney, B. (1978). *The acquisition of morphophonology*. Monographs of the Society for Research in Child Development, 43.
- Murati, Q. (1989). E folmja shqipe e rrethinës së Tuhinit të Kërçovës. *Studime gjuhësore II (Dialektologji)*, Prishtinë, 3-113.
- R Core Team. (2020). *R: A language and environment for statistical computing*. <https://www.R-project.org>
- Schiel, F. (1999). Automatic phonetic transcription of non-prompted speech. *Proceedings of ICPhS 14*, San Francisco, 607-610.
- Shkurtaç, Gj. (2004). *Etnografi e të folurit të shqipes*. SHBLU. Tiranë.
- Stemberger, J. (1993). Vowel dominance in over-regularizations. *Journal of Child Language*, 20, 503-521.
- Tessier, A. (2016). Morpho-Phonological Acquisition. In Lidz, J., Snyder, W. & Patter, J. (Eds.), *The Oxford Handbook of Developmental Linguistics*. Oxford. Oxford University Press.
- Tomas, E., van de Vijver, R., Demuth, K. & Petocz, P. (2017). Acquisition of nominal morphophonological alternations in Russian. *First Language*, 37, 453-74.
- Winkelmann, R., Harrington, J. & Jänsch, K. (2017). EMU-SDMS: Advanced speech database management and analysis in R. *Computer Speech & Language*, 45, 392-410.



Many thanks go to:

Children, teachers, parents, grandparents of selected grade 1 classes at the “Qazim Turdiu” and “Haki Shehu” schools in the Tirana greater area for their kind participation in this study.

This research was funded by the European Research Council (ERC) under the European Union’s Horizon 2020 research and innovation programme (grant agreement No. 742289).