# EJECTIVES IN GEORGIAN. A REAL-TIME MRI ANALYSIS OF VERTICAL LARYNX MOVEMENT

Nato Sulaberidze [1], Erika Brandt [2], Phil Hoole [3], Martin Krämer [4], Jürgen R. Reichenbach [4], Adrian P. Simpson [1]

[1] Friedrich Schiller University Jena, [2] State Criminal Police Office Berlin, [3] Ludwig Maximilian University Munich, [4] Jena University Hospital
nato.sulaberidze@uni-jena.de, erika.brandt@polizei.berlin.de, hoole@phonetik.uni-muenchen.de, martin.kraemer@med.uni-jena.de, juergen.reichenbach@med.uni-jena.de

## ABSTRACT

Vertical larynx movement is often claimed to be responsible for the intraoral pressure increase in ejectives. However, the epiphenomenal production of these sounds in German shows no motivation for such laryngeal involvement and appears to be driven by pulmonic airflow. It is plausible that in a language with phonological ejectives, these stops could also be produced, at least in part, by pulmonic airflow and no larynx involvement. The present analysis aims to investigate this by examining vertical larynx movement in the production of ejectives in Georgian. As part of a larger study, real-time MRI recordings of Georgian speakers are examined. Significant larynx height and movement shape differences between ejectives and their non-ejective congeners (voiceless aspirated and voiced plosives) are observed. Ejectives showed a greater average larynx height than control plosives. Large laryngeal displacement was observed only in the sentence-initial stops, regardless of the stop category.

**Keywords:** Ejectives, Georgian, vertical larynx movement, larynx height, real-time MRI

## 1. INTRODUCTION

The textbook description of an ejective is an obstruent (plosive, fricative) produced with an airstream mechanism involving a closed glottis and an upward movement of the larynx that reduces the volume of the supraglottal cavity. If the velum is raised and there is a complete closure or a stricture of close approximation somewhere in the oral cavity, intraoral pressure increases fuelling the plosive release or the turbulent airflow for a fricative [1]. In German, the release of word-final plosives preceding glottalised vowel-initial syllables typically exhibit the auditory and acoustic characteristics of ejective stops [2]. However, in these epiphenomenal ejectives, arising from the temporal overlap of oral and glottal closure in this context, there is no motivation for any vertical larynx movement [3]. Instead, there is evidence that the air pressure fuelling the release is created by a pulmonic airstream flowing into the closed oral cavity *before* the glottis is closed. With the glottis closed or configured for a creaky voice, the release acquires the auditory and acoustic characteristics of an ejective [2, 4]. Accounting for the production mechanisms behind epiphenomenal ejectives in German has two interesting consequences. First, it provides a plausible path for the emergence of ejectives in a language in which they are not (yet) part of the regular phonological system [5]. Second, although we have just described the production of pulmonically fuelled ejective-sounding plosives in German, there is no reason why languages with ejectives as part of their regular phonological inventory should not also exploit a similar method of production, i.e. use a pulmonic airflow during the stop closure for intraoral pressure build-up before closing the glottis prior to stop release. Indeed, this account has been entertained in an earlier study modelling airflow in ejectives [6].

The present study explores the second of these possibilities by using real-time magnetic resonance imaging (rtMRI) to examine whether there is evidence for vertical larynx movement during the production of ejectives in Georgian, a language in which they make up part of the regular phonological inventory [7].

## 2. METHOD

As part of a larger study comparing the production of ejectives in Georgian, English, and German using dual-channel electrolaryngography and intraoral pressure measurement [3], a subset of 5 Georgian and 7 German speakers were recorded producing target stops in a range of contexts relevant to each language using rtMRI. In the current study, we focus on the analysis of the Georgian data.

### 2.1. Speech material

In Georgian, ejectives stand in regular phonological contrast to voiced and voiceless aspirated stops [7]. To allow for comparison between ejectives and their pulmonic congeners, Georgian speech material comprised words (near-minimal pairs) embedded in a sentential context containing either ejectives or their voiceless aspirated and voiced congeners (controls) at bilabial /p' p b/, alveolar /t' t d/ and velar /k' k g/ places of articulation. During the recordings, speech material was projected onto a screen in the MRI device. Subjects first read a short text and then the short sentences, each presented individually and in randomised order. For the present analysis, only the sentence material was analysed. In the elicited material, vocalic context and the position of the obstruents in the sentence and word were controlled. Target and control stops occurred at three different places: (a) utterance-/sentence-initial (e.g. პალატა /**p'**ala'ta/ 'hospital room' ) and in the second word of the sentence, either in (b) word-initial (სუფთა პანელი /supta **p'**aneli/ 'clean panel') or (c) word-medial position, at the beginning of the second syllable (e.g. წვერის საპარსი /t͡s'veris sa**p'**arsi/ 'beard razor'). Stops occurred in the vocalic context with open and close vowels depending on the sentence context, as follows: Ca & Ci (sentence-initial), aCa & iCi (intervocalic). In total, the sentence material contained 162 target words (3 places of articulation $\times$ 3 manners of articulation $\times$ 3 sentence contexts $\times$ 2 vowel contexts $\times$ 3 elicitations).

### 2.2. Subjects

Five female native speakers of Georgian ($n = 5$, aged 25–30 years, $M = 27(2)$) were recorded. They were students at the University of Jena. One of the subjects had to be excluded because images were found to be off the midsagittal plane obscuring the tracheal airway needed for visually establishing the lower edge of the larynx. The remaining four subjects gave Tbilisi as the place where they had lived the longest.

### 2.3. Experiment Setup

Subjects were recorded at the Jena University Hospital using a clinical 3T whole-body MRI scanner (Magnetom PRISMA fit, Siemens Healthineers, Erlangen, Germany) with the 64-channel head coil supplied by the manufacturer. For sound recording, a Sennheiser MO2000 optical microphone was attached to the head coil at 3 to 4 cm from the subject's mouth. For data acquisition, a single slice 2D radial Fast Low Angle-Shot measurement sequence [8] was used using a golden angle temporal ordering [9] and a radial undersampling factor of 7. Parameters: Echo time: 1.2 ms, fat saturation: off, flip angle: 8°, matrix size: 96 x 96, slice thickness: 6 mm, readouts per fully sampled k-space: 150, readouts per undersampled k-space: 21, repetition time: 2.7 ms, spatial resolution: 2.7 x 2.7 mm$^2$. The slice was aligned in a midsagittal orientation with a field of view size covering the entire head and parts of the neck. Image reconstruction was performed offline using conjugate gradient Sensitivity Encoding (CS-SENSE) [10] and sliding-window reconstruction [11] using a frame-to-frame temporal shift of 19 ms that resulted in a reconstructed frame rate of 53 frames per second.

### 2.4. Data processing

The audio data has been manually synchronised with the MRI images with the Computer Vision Toolbox in Matlab [12]. To remove MRI noise in the recorded signals, the MRI Speech Denoising Toolbox was used [13]. The audio data were automatically segmented and labelled using WebMaus tools [14] and manually adjusted in *Praat* [15]. The acoustically visible release point of the plosives served as the reference time point for the following analysis. The MRI image processing was done with Matlab [12] to extract the vertical position of the lower edge of the larynx over time. Image registration to a reference image was performed to control for possible movements of the speaker's head during the recording session and to obtain a vertical orientation of the trachea immediately below the larynx. Images with neutral laryngeal and head positions during the production of the word-initial [i] sound in the utterance "Sie fahren zur IAA nach Frankfurt", spoken by all subjects at the beginning of the recordings, were chosen as reference. The analysis was based on the lower edge of the laryngeal tissue to avoid the impact of the epiglottis in the signal. A simple edge-detection procedure extracted the vertical position at the transition from air in the trachea (lower intensity) to laryngeal tissue (higher intensity).

A contour of 15 temporally equidistant points representing the position of the lower edge of the larynx was extracted for each of the target

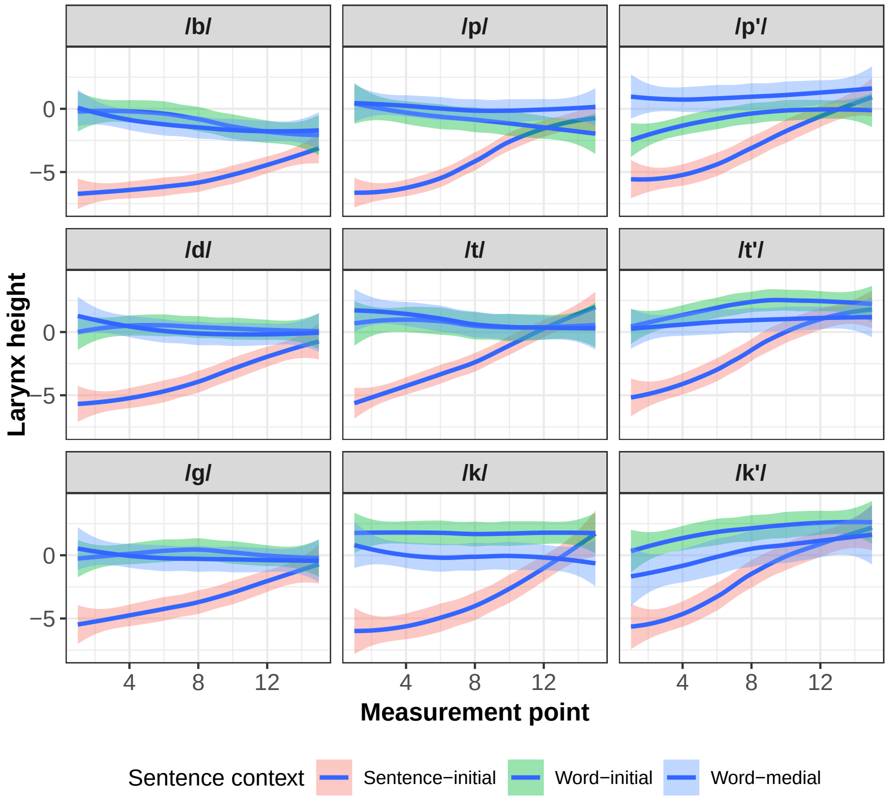


**Figure 1:** Smoothed larynx lower edge contours of target ejective (right), voiced (left) and voiceless aspirated stops (centre) in the three utterance contexts.

and control plosives for a fixed time range of 300 milliseconds centred around the stop release established in the acoustic record. Larynx contours containing outliers related to non-speech movements and measurement errors (NaN values) caused by the larynx not being visible in the recording were removed from the data set.

### 2.5. Data analysis

We used generalized additive mixed-models (GAMMs) for statistical analysis based on the mgcv [16] and itsadug [17] R packages [18]. GAMMs allow statistical analysis of nonlinear regressions. Therefore, they are specifically interesting for phonetic data which is collected in a time series and changes dynamically. We followed the modelling procedure given in [19].

## 3. RESULTS

Figure 1 displays the smoothed larynx lower edge contours of target ejective (right), voiced (left) and voiceless aspirated stops (centre) in the three utterance contexts. Immediately apparent from the plots is the substantial influence of utterance context for all the plosives, with the utterance-initial plosives exhibiting the largest larynx displacement regardless of the airstream mechanism they are being driven by. Although it is difficult to interpret the exact differences in laryngeal height between the ejective, voiceless and voiced plosives from the figure shown, statistical analysis with GAMMs provides more precise information on these differences.

**Listing 1:** Structure of GAMM for larynx height

```
#main effects on larynx height
m = gam(LH ~ Plosive + Context +
#smooth term plosive ordered
      s(MP, by=PlosiveO) +
#smooth term context ordered
      s(MP, by=ContextO) +
#Non-linear random effect
      s(MP, Speaker, bs="fs",m=1),
#Restricted maximum likelihood
      method='REML', rho=rhoval
#Control for auto-correlation
      AR.start=geo$start.event,
#Data specification
      data=geo)
```

The structure of the GAMM for larynx height (LH) is given in listing 1 with the fixed effects plosive category (levels: ejective, voiceless, voiced) and sentence context (levels: sentence-initial, word-initial, word-medial) and additional smooth effects for measurement point (MP) by ordered plosive and ordered context as well as a non-linear random effect which accounts for non-linear differences over measurement point with respect to each speaker. Ordered by-term factors are used to investigate the difference between the individual levels in trajectory shape. Following [19] we control for auto-correlation (see listing 1). Its value is set to a score that is present at a lag of 1.

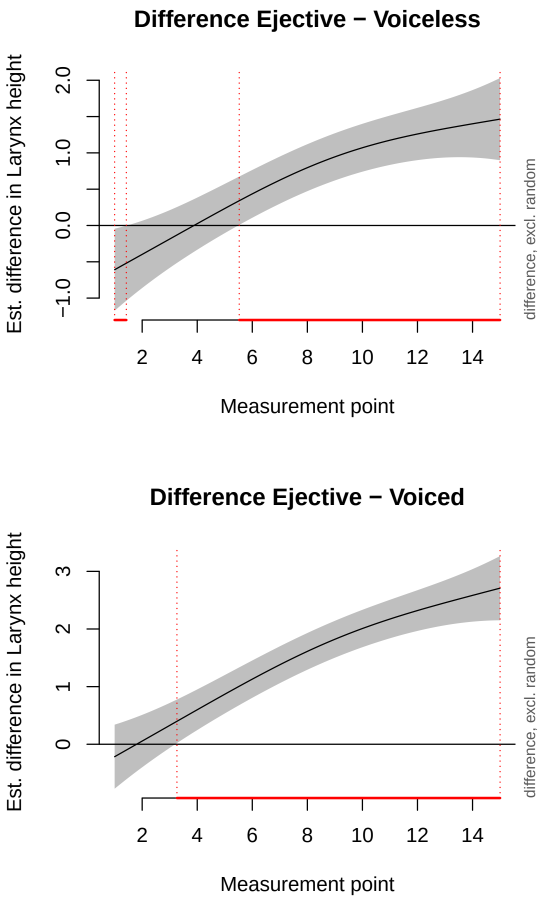


**Figure 2:** Difference smooth in Larynx height for ejectives and voicelsess (top) and ejectives and voiced stops (bottom) with 95% confidence intervals (shaded band). Significant differences are marked in red.

For model comparisons we use Akaike Information Criterion [20]. The nonlinear patterns are modelled with the default smooth type 'thin plate regression spline' [21]. The basis dimension (k) was set to 10 by default for this model. The final model explains 75.4% of data variability for larynx height in Georgian. On average, larynx height for voiced and voiceless Georgian plosives is significantly lower than for Georgian ejectives. Word-initial and word-medial plosives have significantly higher average larynx height than plosives in the sentence-initial position. With regard to the smooth terms, we find significant differences in the shape of the larynx lower edge contours between ejectives and both voiced and voiceless plosives. The trajectory shape for larynx height in sentence-initial plosives is also significantly different from word-medial and word-initial plosives. The random effect for differences over measurement point for each speaker also adds significantly to model performance. Figure 2 shows the estimated smooth differences in the shapes of the vertical movement of the larynx lower edge in ejectives compared to voiceless (top) and voiced stops (bottom). Red marks in the figure indicate the phases where the differences in the shape of the laryngeal trajectory are significant. The figure also distinguishes between the specific differences in larynx height values in the target and control stops, as follows: If the estimated difference values are above zero, then the laryngeal height values in the target stops (ejectives) are higher than in the conditions such as voiced and voiceless aspirated and vice versa, values below zero indicate lower larynx height in target stops compared to voiced or voiceless aspirated plosives respectively.

## 4. DISCUSSION

We have been able to register upward vertical larynx movement in the present study. However, the most remarkable factor contributing to differences in larynx movement is utterance context with the largest upward displacement in sentence-initial position, regardless of the plosive type. In a comparison of the target ejectives with voiceless aspirated and voiced plosives more nuanced differences did arise. In our data, Georgian ejectives show on average significantly higher larynx position than the other stop categories. However, based on the trajectory observations of laryngeal movements during the production of ejective stops compared to the other plosive categories, the suggestion that laryngeal movements in ejectives are pronounced enough to be associated with the increase in intraoral pressure, remains questionable. This can be considered as an indication, that ejective stop releases in many word- and sentence-internal contexts are being driven by pressure built up by a pulmonic rather than a glottalic airstream.

## 5. ACKNOWLEDGEMENTS

This study is funded by the German Research Council (grant: SI743/8-2). We would like to thank the Georgian subjects for their participation.


## 6. REFERENCES


[1] J. Catford, *Fundamental Problems in Phonetics*. Edinburgh: Edinburgh University Press, 1977.

[2] A. P. Simpson, "Ejectives in English and German – linguistic, sociophonetic, interactional, epiphenomenal?" in *Advances in sociophonetics*, C. Celata and S. Calamai, Eds. Amsterdam: Benjamins, 2014, pp. 187–202.

[3] E. Brandt and A. Simpson, "The production of ejectives in German and Georgian," *Journal of Phonetics*, vol. 89, p. 101111, 11 2021.

[4] A. P. Simpson, "Acoustic and auditory correlates of non-pulmonic sound production in German," *Journal of the International Phonetic Association*, vol. 37, no. 2, pp. 173–182, 2007.

[5] J. J. Ohala, "Emergent stops," in *Proc. 4th Seoul International Conf. on Linguistics [SICOL]*. Seoul: Linguistic Society of Korea, 1997, pp. 84–91.

[6] J. C. Kingston, "The phonetics and phonology of the timing of oral and glottal events," Ph.D. dissertation, University of California, Berkeley, 1984.

[7] R. Shosted and V. Chikovani, "Standard Georgian," *Journal of the International Phonetic Association*, vol. 36, pp. 255–264, 2006.

[8] M. Krämer, M. Weirich, K.-H. Herrmann, A. P. Simpson, and J. R. Reichenbach, "Assessing intrinsic velum height in vowels using time-resolved MRI," in *Proceedings 25. Annual Meeting International Society for Magnetic Resonance in Medicine*, vol. 25, Honolulu, Hawaii, 2017, p. 1511. [Online]. Available: http://archive.ismrm.org/2017/1511.html

[9] S. Winkelmann, T. Schaeffter, T. Koehler, H. Eggers, and O. Doessel, "An optimal radial profile order based on the golden ratio for time-resolved MRI," *IEEE Transactions on Medical Imaging*, vol. 26, no. 1, pp. 68–76, 2007.

[10] K. P. Pruessmann, M. Weiger, P. Börnert, and P. Boesiger, "Advances in sensitivity encoding with arbitrary k-space trajectories," *Magnetic Resonance in Medicine*, vol. 46, no. 4, pp. 638–651, 2001. [Online]. Available: https://onlinelibrary.wiley.com/doi/abs/10.1002/mrm.1241

[11] S. J. Riederer, T. Tasciyan, F. Farzaneh, J. N. Lee, R. C. Wright, and R. J. Herfkens, "Mr fluoroscopy: Technical feasibility," *Magnetic Resonance in Medicine*, vol. 8, no. 1, pp. 1–15, 1988. [Online]. Available: https://onlinelibrary.wiley.com/doi/abs/10.1002/mrm.1910080102

[12] MATLAB, *version 9.7.0.1319299 (R2019b)*. Natick, Massachusetts: The MathWorks Inc., 2019.

[13] C. Vaz, "MRI speech denoising toolbox," https://github.com/colinvaz/mri-speech-denoising.

[14] T. Kisler, U. D. Reichel, and F. Schiel, "Multilingual processing of speech via web services," *Computer Speech & Language*, vol. 45, pp. 326-347, 2017.

[15] P. Boersma and D. Weenink, "Praat: doing phonetics by computer [computer program]," 2022. [Online]. Available: https://www.praat.org

[16] S. Wood, *Generalized Additive Models: An Introduction with R*, 2nd ed. Chapman and Hall/CRC, 2017.

[17] J. van Rij, M. Wieling, R. H. Baayen, and H. van Rijn, "itsadug: Interpreting time series and autocorrelated data using GAMMs," 2022, r package version 2.4.1.

[18] R Core Team, *R: A Language and Environment for Statistical Computing*, R Foundation for Statistical Computing, Vienna, Austria, 2022. [Online]. Available: https://www.R-project.org/

[19] M. Wieling, "Analyzing dynamic phonetic data using generalized additive mixed modeling: A tutorial focusing on articulatory differences between l1 and l2 speakers of english," *Journal of Phonetics*, vol. 70, pp. 86–116, 2018. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0095447017301377

[20] H. Akaike, "A new look at the statistical model identification," *IEEE Transactions on Automatic Control*, vol. 19, no. 6, pp. 716–723, 1974.

[21] S. Wood, "Thin plate regression splines," *Journal of the Royal Statistical Society:Series B (Statistical Methodology)*, vol. 65, no. 1, pp. 95–114, 2003.