Larynx raising in English word-final ejective stops: A real-time MRI study

Lavinia Price, Marianne Pouplier, Philip Hoole

Institute for Phonetics and Speech Processing (IPS), LMU Munich
{l.price | pouplier | hoole}@phonetik.uni-muenchen.de

ABSTRACT

This study addresses the nature of production mechanisms underlying English ejectives. Using a corpus of real-time MRI data from 27 native British English speakers, it investigates the extent to which English word-final ejective stops are produced by the traditional 'Catfordian' mechanism with active larynx raising. Overall, 8.5% of word-final stops in the dataset are auditorily identified as ejectives. Analysis of larynx position during the ejectives' closure interval reveals a raised position in only 48.6% of cases, suggesting that more than one mechanism is involved in ejective production. Acoustically, these ejectives are significantly distinct from their plain counterparts in the measures applied. Contextual factors known to promote ejective stop production in English are confirmed in our dataset, in that they appear predominantly as allophones of voiceless velar stops in stressed words.

Keywords: ejectives, rt-MRI, airstream mechanism, speech production

1. INTRODUCTION

Ejectives are traditionally defined as the products of a glottalic airstream caused by raising of the larynx during simultaneous constrictions at the glottis and in the oral cavity [3]. Larynx-raising reduces the supraglottal cavity, thereby increasing the intraoral air pressure (IOP) which results in the auditorily distinct quality of these sounds. An ongoing debate surrounds the hypothesis that ejectives may, however, result from a range of different articulatory mechanisms which may or may not involve larynx raising [1, 9, 10, 14, 15, 21, 23, 25]. This in turn means that both a glottalic or pulmonic airstream initiation may be possible for ejectives. How non-glottalic initiation may come about in detail and what the articulatory-acoustic properties of pulmonic ejectives might be remains, however, unknown. The goal of the current paper is to contribute to this debate by investigating larynx position in English ejective production.

1.1. Non-glottalic ejective initiation

Kingston [9] voices doubts regarding the larynx's efficacy in creating the increase in IOP necessary for the high intensity bursts found in some types of ejectives, and suspects the involvement of additional articulators, e.g., tongue root backing. Most recently, Brandt and Simpson [1] compared articulatory mechanisms of Georgian phonological and German epiphenomenal ejectives using dual-channel EGG. Their results indicate no larynx raising at all during the production of German ejectives that occur epiphenomenally when a final stop is overlapped in time by a following glottal stop. In Georgian, they show that larynx activity is variable and that it depends on the position of the ejective within the word or sentence. Combined with IOP measurements they find these results to support Simpson's theory of a type of 'pulmonic' ejective which does not require larynx raising to be initiated [21]. In addition, the traditional tense vs. lax ejective categories, for which differences in vocal fold stiffness and longitudinal tension are proposed to be the articulatory bases [9, 14], are considered too restrictive. The high intra- and inter-language variation found in an increasing number of studies on the acoustic-phonetic properties of ejectives [7, 8, 15, 22, 24, 25], suggests rather that the phonetic realization of ejective sounds may be much more variable than their textbook description allows for [4, 12, 22]. Evidently, there are fundamental gaps in our knowledge regarding the variability of these sounds, and it is against this background that English ejectives represent a particularly interesting case.

1.2. English ejectives

Ejectives are not originally part of the phonological inventory of the English language. Yet, their prevalence in every-day speech is not only well-established, but evidence points to them being phonetic variables at a "prior stage to a more systematic phonological status" [21] (p.191). They have been described to serve a range of interactional, prosodic, and sociophonetic functions, within the English sound system [5, 16, 17, 21]. Structural contexts are known to influence their realization. They appear exclusively word-finally, preferably at the end of phrases or before pauses [17]. It has been proposed that their diachronic emergence in English has most likely been triggered by an internal language source, yet the details of such an account are currently unclear, as is the nature of how these ejectives are actually produced [21]. These latter points raise questions of broader scientific relevance. The English consonant system is purely pulmonic. What are the circumstances under which sounds of (presumably) glottalic nature develop within this system? What induces an English speaker to actively employ larynx raising during the realization of a stop? In fact, Simpson suspects at least a portion of English ejectives to be produced *without* larynx raising. He suggests that the IOP increase typical for ejectives is caused instead by a pulmonic airstream which continues to emanate through the open glottis after the onset of oral closure. This in turn is, as Simpson suggests, a consequence of the temporal realignment of neighboring glottal and oral articulations as they occur when a glottal stop follows a word-final plosive [21]. Simpson thus provides a pulmonic account of ejective production in English, obviating the puzzle of how glottalic initiation may arise in a purely pulmonic sound system. Yet the extent to which larynx raising actually happens in English ejectives remains poorly understood, due also to the methodological difficulties in observing larynx position. The current study capitalizes on an existing rt-MRI corpus to uncover some of the articulatory and acoustic characteristics of English ejectives. The rt-MRI of a sizable number of speakers offers a unique opportunity to investigate the extent to which ejectives are initiated by vertical larynx movement and determine the acoustic characteristics distinguishing these stops from their pulmonic counterparts.

2. METHODS

2.1. Participants and recordings

This study features rt-MRI data from 27 speakers of Standard Southern British English (13F; mean age=24). Data was collected for an independent study at the Max-Planck-Institute for Biophysical Chemistry (Göttingen, Germany) at 50.05 frames per second with a temporal resolution of 20 ms and an in-plane pixel resolution of 1.4 x 1.4 mm using a 3T MRI system (Magnetom Prisma Fit, Siemens Healthineers, Erlangen, Germany). Synchronized, noise-supressed audio was obtained during the scanning session by means of an Optoacoustics FOMRI III fiber-optic dual-channel microphone (Optoacoustics LTD).

2.2. Materials

The available stimuli consist of 130 lexical items with voiced and voiceless coda stops at three places of articulation (POA), embedded sentence medially into a variety of carrier phrases. Stimuli were produced once per phrasal stress condition (accented vs. deaccented). A total of 5879 tokens were analyzed, 4451 alveolar, 651 bilabial, and 777 velar (items were not balanced for POA, as this corpus stems from an unrelated study on nasalization). Codas were either simplex (/-p, -t, -k, -b, -d, -g/) or /NC/-clusters (/-nt, -nd, -mp, -md, -ŋk, -ŋd/) with a range of preceding vowels and followed by the vowel /ʌ/, e.g., "*Saw **bat** about six*".

The noise-suppressed audio collected during the MRI scanning sessions served as a basis both for stop categorization in the auditory analysis and for the subsequent analysis of acoustic features. The target stop, its closure, and the preceding vowel were manually segmented using Praat [2]. Boundaries marked the interval between the target stop's burst release and the onset of the following /ʌ/.

2.3. Analyses

2.3.1. Auditory categorization

The first question to be addressed concerns the frequency and distribution of ejective realizations. For this, all coda stops in the dataset were auditorily categorized by the first author as either 'ejective' or 'other'. The latter group comprises the subcategories 'pulmonic', 'glottal stop', i.e., the realization of a pure glottal stop with no oral constriction, and 'unclear', i.e., auditorily ambivalent cases.

2.3.2. Acoustic analysis

For this analysis, only tokens from the 'ejective' and 'pulmonic' categories of the 22 speakers who ultimately produced ejectives were retained (n=1622).[1] A range of acoustic measures was selected in order to identify how ejectives may be acoustically distinct from their pulmonic counterparts. Their choice was informed by studies on phonological ejectives in other languages [6, 7, 22, 23, 24, 25]. These measures target the ejective's salient burst quality and the typical interval of acoustic silence caused by the delayed release of the glottal relative to the oral release. *Burst intensity* (dB) was calculated as the mean intensity of the first 40 ms of the burst; *burst intensity difference* (Δintensity) was defined as the difference in mean intensities (dB) between the first 20 ms of the release burst and the following 20 ms interval (Δintensity = mean intensity$_{\text{phase2}}$ – mean intensity$_{\text{phase1}}$). This measure is designed to quantify the steep fall-off in intensity that an ejective burst may have in contrast to a pulmonic stop burst. *Break intensity* was the mean intensity (dB) of the silent interval characteristically following an ejective release, quantified here as starting 40 ms after the burst and ending with the onset of the following /ʌ/.

Closure duration (ms) was the interval between closure onset and release of the target consonant. *Break duration* (ms) was calculated as the difference between the release of the target consonant closure and the onset of the following vowel (/ʌ/).

2.3.3. Articulatory analysis

The articulatory movement of the larynx was observed during ejective, pulmonic voiceless and nasal consonants. Larynx movement was quantified by defining a region of interest (ROI) around the bounds of the larynx and tracing the changes in pixel intensity within that region. Larynx displacement (Δlarynx) was calculated as the difference between two positional values extracted from the larynx trace. These values correspond to the position of the lower larynx's edge at the moment of the stop's closure (P1) and burst release (P2), as seen in Figure 1. Positive Δlarynx values mean a raised larynx position at the release relative to the onset of closure, a negative value a lower larynx position.

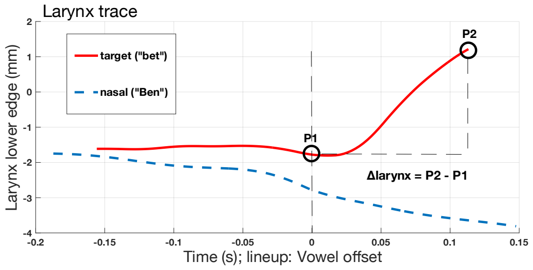

Figure 1: Calculation of larynx displacement using positional values extracted from the ROI's larynx trace. P1=word-final stop closure onset, P2=burst release for plosives/closure offset for nasals.

2.3.4. Statistical analyses

For all continuous articulatory and acoustic measures, linear mixed effects models were conducted using the lmerTest package in R [11, 20]. The five acoustic measures as well as the Δlarynx values were the dependent variables. Duration measures were square root transformed and intensity measures z-score normalized. Stop type (3 levels: ejective, pulmonic voiceless and pulmonic voiced for the acoustic analysis; ejective, pulmonic voiceless and nasal for the articulatory analysis) and POA (3 levels: bilabial, alveolar, velar) were included as fixed factors and a random intercept added for both speaker and lexical item. Pairwise comparisons were conducted using Tukey's post-hoc tests. Significance was evaluated at the $p<.05$ level.

3. RESULTS

3.1. Frequency and structural conditioning

8.5% (n=502) of all word-final stops in the dataset were auditorily identified as ejectives. Five speakers did not produce any.

Place			Voicing		Accent	
bilabial	alveolar	velar	voiceless	voiced	accented	deaccented
11.2%	5.9%	21.4%	12.5%	4.5%	12%	1.2%

Table 1: Percentage of ejectives realized within each phonetic category (for each level of the category 100% is reached adding the non-ejective realizations).

Table 1 shows the influence of phonetic categories on the realization of ejectives. The percentages confirm previous findings that ejectives in English predominantly appear as allophones of originally voiceless, velar stops in accented words [17]. The fact that ejectives replace voiced stops at all, however, is worth noting.

3.2. Acoustic properties of English ejectives

Figure 2 gives the results for the acoustic measures. Following [22], both voiceless and voiced pulmonic stops were considered in the acoustic comparison with ejectives. For all acoustic measures but break duration, main effects of stop type and POA were significant at $p<0.001$ (for burst intensity and closure duration stop type was significant at $p<0.01$). Significant interactions between POA and stop type were found for all measures but closure duration ($p<0.001$).

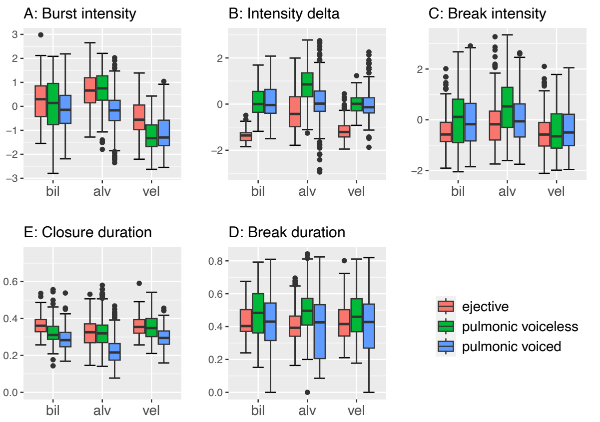

Figure 2: Acoustic measures (A-C z-normalized, D and E square root transformed) by stop type (ejective, pulmonic voiceless, pulmonic voiced) and place of articulation.

Tukey post-hoc tests show that Δintensity is the only measure which significantly distinguishes ejectives from both pulmonic voicing types at all three POAs ($p<0.001$), as illustrated in Figure 2B. A significant difference in burst intensity (Figure 2A) is found for ejectives and both pulmonic stop types at the velar POA (/k’/vs./k/=$p<0.001$ and /k’/vs./g/=$p<0.05$) as well as between alveolar ejectives and pulmonic voiced stops (/t’/vs./d/=$p<0.001$). Break intensity (Figure 2C) differs significantly for ejectives and pulmonic voiceless stops at both the bilabial (/p’/vs./p/=$p<0.001$) and alveolar (/t’/vs./t/=$p<0.001$) POA. Overall, these results suggest that the dynamics of burst intensity are the most consistent acoustic correlate of English ejective production.

3.3. Larynx raising in English ejectives

Figure 3 gives the results for the larynx position measure. It shows a three-way comparison between Δlarynx values calculated for all ejective, pulmonic voiceless and nasal stop tokens in the dataset. Significant main effects of stop type (F[2,30.3] =14.3, p<0.001) and POA (F[2,72]=7.2,p<0.01) were found on Δlarynx values, as well as a significant stop type by POA interaction (F[4,300.1]=17.3,p<0.001). Larynx displacement is significantly different between velar ejectives and both other velar stop types ($p<0.001$) as well as between bilabial ejectives and their pulmonic counterparts ($p<0.05$), however, not between any of the three alveolar stop types.

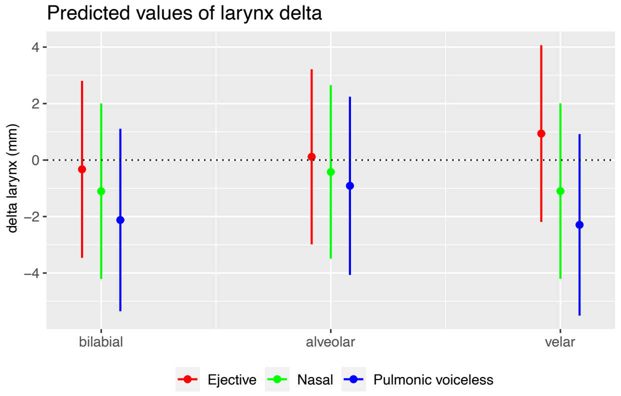

Figure 3: Marginal effects of the three stop types (ejective, nasal and pulmonic voiceless) on Δlarynx values. Positive values indicate raising, negative larynx lowering.

Figure 3 shows a uniform pattern in the predicted Δlarynx values, dependent on the stop type. Across all POAs, the highest values are predicted for ejectives, these are followed by nasals and are lowest for pulmonic voiceless stops. Homogenous larynx behaviour, i.e., larynx lowering, is found for both pulmonic voiceless and nasal stops at all POAs. This is in agreement with previous findings on larynx lowering for nasals in English [19]. Importantly, there is no consistent larynx raising for ejectives across POAs. According to these results, only velar ejective initiation occurs by clear larynx raising in the sense of Catford [3].

4. DISCUSSION AND CONCLUSION

The goal of our study was to investigate larynx raising in English ejectives and to examine their acoustic profile. Articulatorily, we show that ejective initiation is variable in English, they may occur with or without larynx raising. The likelihood of traditional 'Catfordian' initiation, however, is strongly influenced by POA. Velar ejectives are the only ones clearly associated with larynx raising. Our acoustic measures underscore that the most consistent ejective characteristic across POA is burst intensity difference. The two acoustic measures directly targeting burst qualities (burst intensity and Δintensity) are particularly effective at characterizing velar ejectives. This agrees with the impression that /k’/ is the most auditorily salient of the English ejective stops. Combined with the articulatory results, we infer that a larynx raising gesture has a direct effect on ejective burst intensity.

Remarkably, this study finds clear evidence of glottalic initiation in English ejective production. At the same time, it is clear from the results that the Catfordian mechanism does not hold for all of the ejectives in our corpus, and glottalic and (presumably) pulmonic initiation seem to occur side-by-side. It is unclear whether this variation in initiation found here is due to the emergent status of ejectives in English (i.e., a sound change in progress [18]), or whether this may likewise be found among phonological ejectives of other languages. The current study lends further empirical support for the possibility that larynx raising is not the only production mechanism for ejectives [1, 9, 10, 21, 25], even though the exact nature of these 'other' ejective variants remains to be uncovered. It stands to be tested whether these ejective variants are indeed products of a pulmonic airstream mechanism as described by Simpson [1, 21]. Another possibility to be considered is that alternative articulators are acting for the larynx in its role in cavity reduction and IOP increase. It remains to be explored what these articulators may be, other than Kingston’s suggestion of tongue root backing [9, 10]. It is conceivable that variations of ejectives, including those outlined above, may co-exist within and across languages to form an articulatory range of ejective production, with pulmonic and glottalic initiations representing two variants among possibly several others.

5. ACKNOWLEDGEMENTS

Work supported by ERC Horizon 2020 grant Interaccent to J. Harrington.

6. REFERENCES

[1] Brandt, E., Simpson, A.P. 2021. The production of ejectives in German and Georgian. *Journal of Phonetics* 89, 101111.

[2] Boersma, P., Weenink, D. 2019. Praat: Doing phonetics by computer [Computer program], version 6.0.50.

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

[4] Fallon, P. D. 2002. *The Synchronic and Diachronic Phonology of Ejectives*. New York, NY: Routledge.

[5] Gordeeva, O., Scobbie, J. 2013. A phonetically versatile contrast: Pulmonic and glottalic voicelessness in Scottish English obstruents and voice quality. *Journal of the International Phonetic Association* 43(3), 249–271.

[6] Gordon, M., Applebaum, A. 2006. Phonetic structures of Turkish Kabardian. *Journal of the International Phonetic Association* 36(2), 159-186.

[7] Grawunder, S., Simpson, A.P., Khalilov, M., 2010. Phonetic characteristics of ejectives – samples from Caucasian languages. In: Fuchs, S., Toda, M. (eds), *Turbulent sounds. An interdisciplinary guide.* De Gruyter Mouton, 209–244.

[8] Ham, S. 2007. Tsilhqut'in ejectives. *2007 Annual Conference of the Canadian Linguistic Association,* University of Victoria.

[9] Kingston, J.C. 1985. *The phonetics and phonology of the timing of oral and glottal events*. Ph.D. thesis, University of California, Berkeley.

[10] Kingston, J.C. 2005. The phonetics of Athabaskan tonogenesis. In: Hargus, S., Rice, K. (eds), *Athabaskan prosody*. Amsterdam: John Benjamins, 137–184.

[11] Kunzetsova, A., Brockhoff, P., & Christensen, R. (2017). lmerTest package: tests in linear mixed effect models. *J Stat Softw* 82, 1-26.

[12] Ladefoged, P., 1973. The features of the larynx. *Journal of Phonetics* 1, 73–83.

[13] Ladefoged, P., Maddieson, I. 1996. *The Sounds of the World's Languages*. Oxford: Blackwell.

[14] Lindau, M. 1982. Phonetic differences in glottalic consonants. *Journal of the Acoustical Society of America* 71, S1. 22.

[15] Lindau, M. 1984. Phonetic differences in glottalic consonants. *Journal of Phonetics* 12, 147–155.

[16] McCarthy, O., Stuart-Smith, J. 2013. Ejectives in Scottish English: A social perspective. *Journal of the International Phonetic Association* 43(3), 273–298.

[17] Ogden, R. 2009. *An introduction to English phonetics*. Edinburgh: Edinburgh University Press.

[18] Ohala, J. J., 1989. Sound change is drawn from a pool of synchronic variation. *Language change: Contributions to the study of its causes*, 173-198.

[19] Riordan, C.J., 1980. Larynx height during English stop consonants. *Journal of Phonetics* 8, 353–360.

[20] RStudio Team (2020). *RStudio: Integrated Development Environment for R.* RStudio, PBC, Boston, MA URL http://www.rstudio.com/.

[21] Simpson, A.P. 2014. Ejectives in English and German – linguistic, sociophonetic, interactional, epiphenomenal? In: Celata, C., Calamai, S. (eds), *Advances in Sociophonetics*. Amsterdam: John Benjamins, 187–202.

[22] Vicenik, C. 2010. An acoustic study of Georgian stop consonants. *Journal of the International Phonetic Association* 40(1), 59–92.

[23] Warner, N. 1996. Acoustic characteristics of ejectives in Ingush. *Proceedings of the 4th International Conference on Spoken Language Processing*, Philadelphia, 1525–1528.

[24] Wysocki, T.M. 2004. *Acoustic analysis of Georgian stop consonants and stop clusters*. Ph.D. thesis, University of Chicago, Chicago.

[25] Wright, R., Hargus, S., Davis, K. 2002. On the categorization of ejectives: Data from Witsuwit'en. *Journal of the International Phonetic Association* 32, 43–77.

---

[1] The 'unclear' and 'glottal stop' categories were considered irrelevant for the current research question and were removed from further analysis.