PHARYNGEALIZATION OF EAST THURINGIAN POSTVOCALIC /r/: ARTICULATION, ACOUSTICS AND TEMPORAL EXTENT

Christina Otto & Adrian P. Simpson

Institute of German Linguistics, Friedrich Schiller University Jena
otto.christina@uni-jena.de; adrian.simpson@uni-jena.de

ABSTRACT

This study investigates the acoustic and articulatory influence of postvocalic /r/ on the realization of vowels in an east central german variety. A selection of /r/ and /r/-less word pairs containing the short vowels /ʏ ɛ a ɔ ʊ/ of seven male speakers were recorded using Ultrasound Tongue Imanging (UTI). Systematic differences were found both in lingual configurations as well as spectral measures (F1-F3) consistent with pharyngeal constriction. Especially the articulatory analysis of a greater portion of the word indicates consequences of /r/ having greater temporal extent showing /r/ to be a feature of the entire syllable rather than of a single segment.

Keywords: Ultrasound, East Thuringian, /r/, Articulation, Pharyngealization

1. INTRODUCTION

The reduction of consonantal /r/ in coda position to a pharyngealized vowel is described as a distinct and salient feature of east central German varieties [6, 14, 23] but also in south western varieties such as Swabian [12]. As Khan [14] shows for Upper Saxon, /r/ surfaces in this position as pharyngealization of the preceeding vowel and also on adjacent consonantal material. Due to the temporal overlap this results in more monophthongal pharyngeal qualities in combination with short vowels (e.g. [ʃpɔːˤt] *Sport* 'sport') [22].

The dark or 'muffled' quality of pharyngealized segments is the acoustic consequence of the articulatory configuration of the vocal tract. The narrowing in the pharyngeal passage is either produced by contracting the constrictor muscles or by retracting the tongue root [16]. The enlargement of the oral cavity results in a lowered F2 while the smaller pharyngeal cavity results in a raised F1 [24]. Similar patterns of /r/ have been described for Scottish English using Ultrasound Tongue Imaging (UTI) [18, 17]. They show vowel retraction and pharyngealization in derhoticized variants.

While pharyngealization has been studied extensively e.g. in varietes of Arabic [2, 1, 13, 10], articulatory or acoustic studies on east German varieties are scarce [11]. An earlier analyis of the UTI data presented here was able to confirm the influence of /r/ on the tongue contour configuration as both retracted and lowered and on acoustic output (F1 raised, F2 lowered) in target words with postvocalic /r/ in comparison to target words without /r/ [19]. East Thuringian belongs to the east central German dialect group Thuringia-Upper Saxon and is spoken in the east of Thuringia and south western Saxony-Anhalt and is mostly a transition zone to Upper Saxon [23].

While the initial study concentrated on a qualitative analysis and visual description of tongue contours, the aim of the present study is to quantify the differences between tongue contours of vowels with and without following /r/. The research questions are: (1) How does the tongue configuration differ in word pairs with and without /r/? (2) What is the temporal extent of the coarticulatory influence of /r/?

2. MATERIALS AND METHODS

2.1. Data collection

Seven male speakers aged between 26 and 28 from the south of Saxony-Anhalt were recorded. They all belong to the same circle of friends, along with the first author, which, as far as possible, ensured an informal situation during recording.

Ultrasound data of the mid-sagittal contour of the tonge was captured using a transducer type 35C20EA attached to a Mindray DP2200 Ultrasound system. The transducer frequency was 2,5 MHz with a depth setting of 12.9 cm, an internal frame rate of 58 fps and a field of view of 90°. After de-interlacing the video frame rate was approximately 60 fps [28]. The acoustic signal was recorded using a condenser microphone (AKG C1000S) and processed with Articulate Assistant Advanced software (AAA, [5]) along with the video signal. Accurate synchronization of the video and the audio signal was achieved using a synchronization impulse produced by the SyncBrightUp unit [4]. An accurate estimate of the video frame rate of the ultrasound machine was calculated in order to match the audio signal correctly. The ultrasound probe was kept in a relatively stable position using a purpose-built headset (for details see [21]), however, no additional video signal or bite plane was captured to verify that there were no movement of the probe during the recording. Palate traces were recorded only once at the beginning of the recording session whilst the subject swallowed mineral water [27].

Data presented here is taken from seven word pairs containing the short vowels /ʏ ɛ a ɔ ʊ/ and differing in the presence or absence of coda /r/. Targets were embedded in meaningful sentences. The sentence frame for each pair of words was kept the same but varied between different pairs of words, e.g. *Sie ist nach Born/Bonn gefahren* ('She went to Born/Bonn'), *Sie hat den Metz/März geliebt* ('She loved the mason/March'). The decision against a wordlist and pseudo words which would have allowed for better control of coarticulatory influence of surrounding consonantal material was made to avoid making the target the main focus of interest, to enhance naturalness and to evoke more vernacular speech.

Stimuli were presented on a screen in a randomized order using the facility provided in AAA. Each block was repeated five times in the same order. Some tokens had to be omitted because no synchronization impulse was found in the recording.

2.2. Data analysis

Target words were annotated and segmented in PRAAT [8] and imported back into AAA for further processing and analysis. To capture the temporal extent of the influence of /r/, five normalized time points in the vocalic portion of the word were annotated (beginning (V1), after 25% (V2), after 50% (V3), after 75% (V4) and at the end of the vowel (V5)). In addition, three time points during the following consonant or consonant cluster were annotated. Since the syllable structure of the target words vary, the structural meaning of the time points changes. Nevertheless, the three time points capture the temporal extent of the consonant. For the nasal /n/, lateral /l/ and the fricative /s/ three points (25% (C1), 50% (C2) and 75% after the onset (C3)) were selected. For the plosives /k t/ and the affricate /ts/ two points during the closure (after one third and after two thirds) and one in the middle of the release or in the middle of the following fricative were selected.

Splines were fitted to the visible surface of the tongue in the eight frames closest to these eight time points by superimposing a fan grid consisting of 42 radial axes on each frame. The fitting process was done semi-automatically using an edge-detection algorithm implemented in the AAA software which allowed for manual correction of the splines where the algorithm failed or no tongue contour was visible e.g. due to the shadow of the hyoid bone.

Splines were then exported to a workspace for analysis. An average tongue-surface spline was created for each word and each time point over the five repetitions. In order to quantify the differences in tongue configuration of words with and without /r/ throughout the word, pairwise comparisons were made between the mean spline of the /r/ and /r/-less word pairs. The distance between each pair of average splines was calculated by measuring along the 42 radial axes where the two average splines intersected the fan line. Root mean square (RMS) distances for each average spline pair for each time point were calculated from these interspline distance measures (see [18] and [25] for more details on this method).

The first three formant values were estimated at the annotated time points during the vowel (V1, V2, V3, V4, V5) with PRAAT [8]. For statistical analysis linear mixed models (LMMs), as implemented in the lme4 package [7], were run in R [20]. To find the model with the best fit to the data, comparisons were done in a stepwise fashion by adding more factors (or interactions) in each step to the model and comparing this model with the reduced model without the particular factor (or interaction) in question. We then decided on keeping a factor/interaction depending on the result of the likelihood ratio test (p-value < .05).

3. RESULTS

3.1. Articulatory analysis

In order to abstract away from a certain vowel category an average spline was created for each time point over all vowel categories for the /r/ and /r/-less words. By doing so, all coarticulatory information of different vowels, onsets and following consonants is removed. Since the only differing aspect in these comparisons is the presence or absence of /r/, average splines can be interpreted in terms of the impact of /r/ on tongue configuration over several time points throughout the word allowing comparisons of different time points over word tokens and more importantly over multiple speakers. Since the size, position and rotation of the probe varies for different speakers, splines are normally not directly comparable over multiple speakers.

Figure 1: Average tongue splines for /r/ (dark grey) and /r/-less (light grey) vowels for all eight time points for subject RR with standard deviation indicated by the dotted lines. Palate trace is indicated in red and significant interspline distances between the two splines (two-tailed t-test, $p < .05$) are shown by the thicker spokes on 42 fan lines.

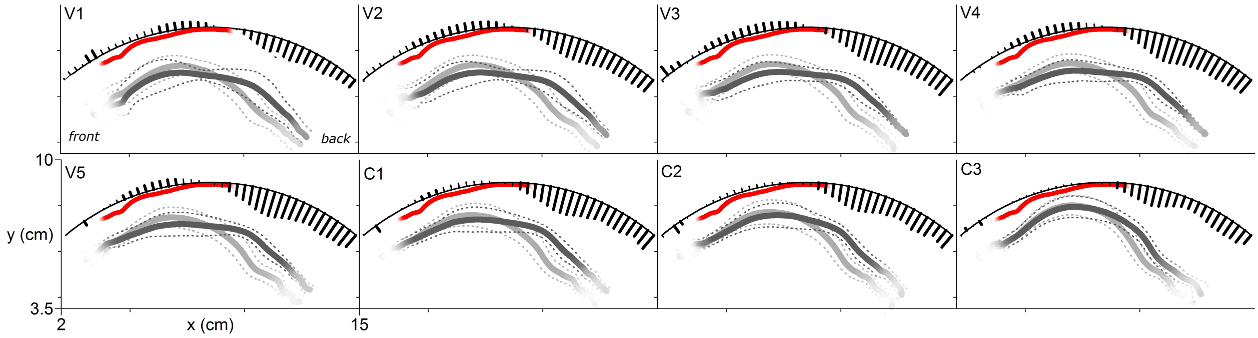


Figure 1 shows the two average splines (thick lines) with standard deviation (thinner dotted lines) of one speaker (RR) at the eight time points. The average splines for the /r/-vowels (dark grey line) are considerably retracted and the shape of the tongue contour is flatter and higher towards the back. The spokes superimposed represent the 42 fanlines and show the difference between the average /r/ and the /r/-less splines. Thicker spokes indicate a significant difference based on a two-tailed t-test with a significance level of $p < .05$. This gives a visual indication of the location of the difference between the two splines over time. The difference in tongue retraction is evident throughout the vowel and the following consonant but decreases – although still significant – at the latest time point in the consonant. The difference in tongue hight, in contrast, is only significant during the vowel and has its peak at the mid point (V3).

The RMS over all speakers is shown in figure 2. A LMM with the RMS distance as dependent variable, timepoint as fixed factor and speaker as random factor shows no significant difference between the time points during the vowel but a significant difference for the time points in the consonant. However, the likelihood ratio test revealed that the LMM with the best fit to the data includes the fixed factors **time point** and **word** as well as an interaction of **time point** and **vowel** ($\chi^2(28) = 104.38$, $p < .001$). This effect can be seen in figure 2 which shows the RMS distances over all speakers for the different time points and word pairs. The interaction of vowel and time point shows up in the *Pack/Park*-pair where the peak of distance is not at the midpoint of the vowel, but rather at the beginning of the vowel. This might be a consequence of the labeling process which was done entirely with reference to the acoustic output and not on articulatory gestures. The vocalic part was labeled from the onset of voicing. In most cases of *Park*, the tongue configuration was already established during the bilabial fortis plosive in the onset and the tongue was already raising for the velar closure. Therefore the difference distribution here might be shifted. In addition, note the difference in the two word pairs containing /ɔ/. Even though /bɔn/ and /ʃpɔt/ have similar underlying phonological structures (bilabial stop in the onset and an alveolar coda consonant), their /r/ vs. /r/-less distance is significantly different. In the pair *Bonn/Born* the difference is larger than in the *Spott/Sport* pair. This might point to a difference in tongue configuration or coarticulatory resistance of German alveolars /n/ and /t/. Furthermore, there is a clear difference for the *Pelle/Perle* pair in three consonantal time points. The influence of /r/ extends over the lateral [l] and remains stable during the consonant. This supports the findings of the earlier study, where acoustic measures (raised F1 and lowered F2) showed that the influence of pharyngealization extended at least into the onset of the following syllable [19].

Figure 2: RMS distances over seven speakers seperated for word pairs and time points during the vowel in light grey and the following consonant in dark grey.

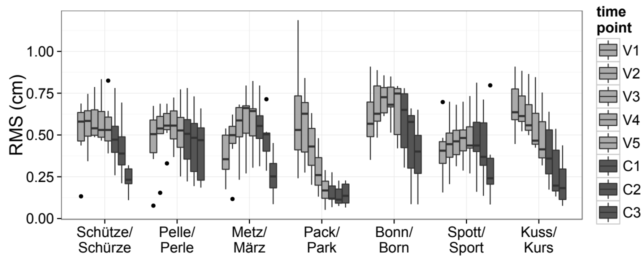
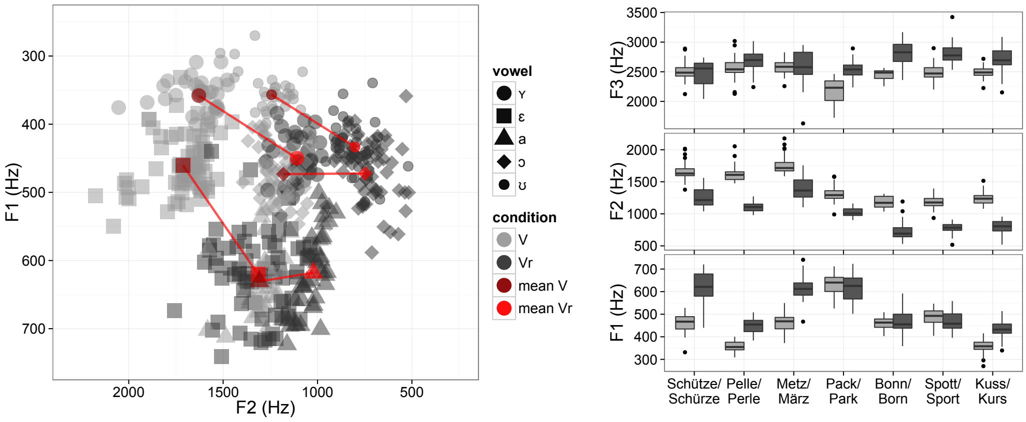


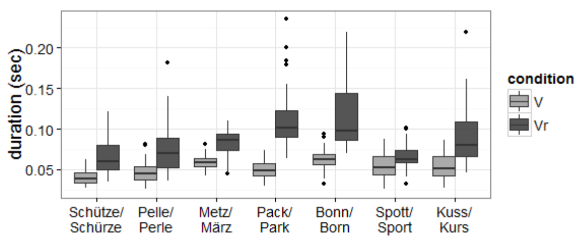
Figure 3: *Left graph:* Formant values F1 and F2 of /r/ vowel in dark grey and /r/-less vowel in lighter grey for all seven speakers. Mean values are in red. *Right graph:* F1, F2 and F3 values seperated for different wordpairs.



3.2. Acoustic analysis

The left graph in figure 3 shows differences in F1 and F2 at the midpoint of the vowel (V3) which can be interpreted with respect to the tongue configurations. /r/-vowels are positioned further back and lower. However, the mean values in red show that the influence of /r/ varies between vowels. A LMM was run over all data with F1, F2, and F3 respectively as dependent variables, with condition as fixed factor and speaker and repetition as random factors. Likelihood ratio tests revealed that the LMM with the best fit to the data for all three formants does not include the interaction with the vowel but with the word pair. The right graph in figure 3 shows this interaction. While for F2 all word pairs differ significantly between the /r/- and /r/-less condition ($\chi^2(4) = 51.96$, $p < .001$), for F1 no significant difference can be found for /ɔ/ and /a/. In addition, the difference between *Pelle/Perle* is smaller than in *Metz/März* ($\chi^2(4) = 10.05$, $p < .05$). The different behaviour for /ɔ/ and /a/ with respect to F1 might be a consequence of preserving contrast between the vowel phonemes, if we take duration (figure 4) into account. The difference between /r/ and /r/-less low vowel /a/ is not realized by lowering the tongue even further but only by decreased F2 (indicating tongue backing) and increased duration. The same is true for *Bonn/Born* and *Spott/Sport*, but with a much smaller difference in duration for *Spott/Sport* than for *Bonn/Born*.

Figure 4: Duration of /r/ (dark grey) and /r/-less vowels (light grey).



The rise in F3 for the low vowel /a/ and the back rounded vowels /ɔ/ and /ʊ/ ($\chi^2(4) = 13.2$, $p < .05$) is in line with Tamimi [1] on Moroccan and Jordanian Arabic whereas Ladefoged and Maddieson consider a lowered F3 as marked feature of pharyngealized vowels [15]. Aralova et al. [3] found higher F3 for vowels with retracted tongue root in a dialect of Even but only for back rounded vowels. This seems consistent with our data even though the reasons for this restriction are not clear and it raises the question on articulatory modelling of the acoustic output.

4. SUMMARY

Tongue configurations for targets with /r/ differ from targets without /r/ with respect to overall shape of the tongue. /r/-vowels tend to be articulated with a flatter tonge body and a retracted tongue root. Retraction seems to be more persistent over time, spreading to adjacent consonantal material. Articulatory analysis of a greater portion of the word has shown that the articulatory correlates of /r/ should be treated as a feature of the entire syllable (and possibly beyond), rather than being merely restricted to a single segment [9, 26].

Acoustic analysis covering duration and spectral measures (F1-F3) support these articulatory findings but the temporal extent in terms of acoustic measures still needs to be carried out. Furthermore, the influence of pharyngealized vowels on F3 needs to be studied more systematically in order to better understand the underlying processes of the acoustic consequences of articulatory settings.

5. REFERENCES

[1] Al-Tamimi, J. 2009. Effect of pharyngealisation on vowels revisited: Static and Dynamic analyses of vowels in Moroccan and Jordanian Arabic. *International Workshop on Pharyngeals & Pharyngealisation*.

[2] Almbark, R. 2009. A Sociophonetic Study of Pharyngealisation in Syrian Arabic differences between Damascene and Aleppian dialects: Gender and regional differences between Damascene and Aleppian dialects. *International Workshop on Pharyngeals & Pharyngealisation*.

[3] Aralova, N., Grawunder, S., Winter, B. 2011. The acoustic correlates of tounge root vowel harmony in Even. *Proceedings of the 17th International Congress of Phonetic Science* 240–244.

[4] Articulate Instruments Ltd, 2010. *SynchBrightUp Users Manual: Revision 1.10*. Edinburgh, UK: Articulate Instruments Ltd.

[5] Articulate Instruments Ltd, 2014. *Articulate Assistant Advanced User Guide: Version 2.15*. Edinburgh, UK: Articulate Instruments Ltd.

[6] Auer, P., Barden, B., Großkopf, B. 1993. Dialektwandel und sprachliche Anpassung bei 'Übersiedlern' und 'Übersiedlerinnen' aus Sachsen. *Deutsche Sprache* 21(1), 80–87.

[7] Bates, D., Maechler, M., Bolker, B., Walker, S. 2014. lme4:Linear mixed-effects models using Eigen and S4 classes: R package version 1.1-7.

[8] Boersma, P., Weenink, D. 2013. Praat: Doing phonetics by computer. Version 5.3.45.

[9] Carter, P., Local, J. 2007. F2 variation in Newcastle and Leeds English liquid systems. *Journal of the International Phonetic Association* 37(2), 183–199.

[10] Embarki, M., Yeou, M., Guilleminot, C., Al Maqtari, S. 2007. An acoustic study of coarticulation in modern standard Arabic and dialectal Arabic: Pharyngealized vs. non-pharyngealized articulations. *Proceedings of the 16th International Congress of Phonetic Science*.

[11] Grawunder, S. 2006. Charakterisierung des R-Lautes im Nordostthüringischen: Ein Werkstattbericht. *19. Sprechwissenschaftliche Fachtagung 10.06.2006*.

[12] Hiller, M. 1995. Regressive Pharyngalisierung im Stuttgarter Schwäbischen als C-V-Interaktion. *Linguistische Berichte* 155, 33–64.

[13] Jongman, A., Herd, W., Al-Masri, M., Sereno, J., Combest, S. 2011. Acoustics and perception of emphasis in Urban Jordanian Arabic. *Journal of Phonetics* 39(1), 85–95.

[14] Khan, S. u. D., Weise, C. 2013. Upper Saxon (Chemnitz dialect). *Journal of the International Phonetic Association* 43(2), 231–241.

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

[16] Laver, J. 1994. *Principles of phonetics*. New York and Cambridge: Cambridge University Press.

[17] Lawson, E., Scobbie, J. M., Stuart-Smith, J. 2011. The social stratification of tongue shape for postvocalic r in Scottish English. *Journal of Sociolinguistics* 15(2), 256–268.

[18] Lawson, E., Scobbie, J. M., Stuart-Smith, J. 2013. Bunched /r/ promotes vowel merger to schwar: An ultrasound tongue imaging study of Scottish sociophonetic variation. *Journal of Phonetics* 41(3-4), 198–210.

[19] Otto, C., Weirich, M., Simpson, A. P. 2014. Spectral and lingual correlates of /r/ in East Thuringia. In: *Proceedings of the 10th ISSP*. Köln: 293–296.

[20] R Core Team, 2014. R: A language and environment for statistical computing: Version 3.1.2.

[21] Scobbie, J. M., Wrench, A. A., van der Linden, Marietta, 2008. Head-probe stabilisation in ultrasound tongue imaging using a headset to permit natural head movement. In: Sock, R., Fuchs, Susanne. and Laprie, Yves, , (eds), *Proceedings of the 8th ISSP*. Strasbourg: INRIA 373–376.

[22] Simpson, A. P. 1998. *Phonetische Datenbanken des Deutschen in der empirischen Sprachforschung und der phonologischen Theoriebildung*. Kiel: Arbeitsberichte des Insituts für Phonetik und digitale Sprachverarbeitung der Universität Kiel 33.

[23] Spangenberg, K. 2013 [1989]. Thuringian. In: Russ, C., (ed), *The Dialects of Modern German*. Hoboken: Taylor and Francis 265–289.

[24] Stevens, K. N. 1998. *Acoustic phonetics*. Cambridge, Mass: MIT Press.

[25] Stone, M. 2005. A guide to analysing tongue motion from ultrasound images. *Clinical Linguistics & Phonetics* 19(6-7), 455–501.

[26] Stuart-Smith, J. 2007. A sociophonetic investigation of postvocalic /r/ in Glaswegian adolescents. *Proceedings of the 16th International Congress of Phonetic Science* 1449–1452.

[27] Wrench, A. A., Cleland, J., Scobbie, J. M. 2011. An ultrasound protocol for comparing Tongue contours: upright vs. supine. *Proceedings of the 17th International Congress of Phonetic Science* 2161–2164.

[28] Wrench, A. A., Scobbie, J. M. 2006. Spatio-temporal inaccuracies of video-based ultrasound images of the tongue. *Proceedings of the 7th ISSP*.