To appear in Interspeech 2015, Dresden

Durational Characteristics and Timing Patterns of Russian Onset Clusters at Two Speaking Rates

Marianne Pouplier[1], *Stefania Marin*[2], *Alexei Kochetov*[3]

[1, 2] Institute of Phonetics and Speech Processing, LMU Munich, Germany
[3] Department of Linguistics, University of Toronto, Canada
pouplier|marin@phonetik.uni-muenchen.de, al.kochetov@utoronto.ca

Abstract

This study presents articulatory data on the durational and timing characteristics of Russian onset clusters and their change as a function of speaking rate. While there is increasing evidence that languages differ systematically in their consonant-to-consonant timing, little is known on whether this difference also entails different implementations of speaking rate changes. Russian contrasts with languages like English or German in that it has little overlap between onset consonants. Relatedly, stop consonants are obligatorily released. We investigate whether these global timing characteristics have implications for the implementation of speech rate changes. We hypothesize that Russian onset timing may vary little as a function of speaking rate, with rate affecting consonant duration rather than C1-C2 timing. Also a cluster's sonority profile (e.g. /bl-/; /lb-/) may factor into the implementation of speech rate changes. Results show, contrary to expectation, that both duration and timing of the consonants in a cluster are subject to change. However, there was a less of a rate effect for clusters with falling sonority, pointing towards their lesser flexibility in timing. Our results also reveal significant differences in the durational control of C1 and C2, challenging current models of durational organization of consonant clusters.

Index Terms: Russian, cluster, speech rate, duration, articulatory timing, overlap

1. Introduction

We present articulatory data on the durational characteristics and coarticulatory patterns in Russian onset clusters and their variation under changes in speaking rate. This work stands in the context of our increasing knowledge of systematic cross-linguistic differences in terms of temporal overlap and consonant-consonant timing. While languages like English, German or French show a relatively high degree of overlap between sequences of consonants, other languages like Russian have been reported to time their consonants far apart [1-5] which may in certain circumstances give rise to a transitional schwa between consonants [6]. Byrd & Tan [7] have found previously for American English cross-word consonant clusters that speaking rate changes are implemented through both changes in maximum constriction lag and durational changes for each member of the C1#C2 cluster. Because they collected EPG data, their evaluation of overlap could only consider onset of tongue-palate contact (essentially the onset of closure) as well as time points of maximum closure. Overlap in terms of when movement into the constriction of C2 begins could not be considered which, to foreshadow one of our results, may render a very different picture on the temporal organization of clusters. Since English and Russian have been shown to differ fundamentally in their consonant timing, speaking rate changes may be implemented differently in the two languages. Pouplier and Beňuš [8] have argued in the context of Slovak syllabic consonants that the range of consonant clusters permitted in a language is related to the degree of overlap between consonants, with more cluster types being attested in languages where consonants in a cluster are timed further apart, i.e. they show greater consonant-consonant lags. From a typological viewpoint, Russian onset cluster phonotactics are unusual, since Russian not only allows for the cross-linguistically most common sonority-rising clusters like /bl-/, but also for sonority plateau (/tk-/) and even sonority falling sonorant-obstruent clusters like /lb-/. Stops in Russian onset clusters are obligatorily released [1-3]. If indeed a certain timing pattern is necessary (for either perceptual or production reasons) to allow for a certain range of phonotactic patterns, we may ask how flexible these timing patterns are under changes in speaking rate. Moreover, since Russian allows for a handful of clusters to combine consonants in either order (e.g., both /bl-/ and /lb-/ are licit clusters), position effects in the cluster can to some degree be separated from segment-specific effects, which would not be possible in most other languages.

Other than the study of Byrd & Tan [7], there is very little work on how speech rate changes affect consonant timing in a cluster. The articulatory implementation of speaking rate has mostly been investigated for consonants and vowels part of repeated CV(C) utterances [9-15] (see [16] for an articulatory study of Russian cross-word clusters). Consonantal duration patterns for Russian have previously been investigated acoustically in interaction with word-edge structure, testing the extent to which duration serves as a listener-oriented cue to boundary location [2, 17]. Redford [17] found that prevocalic stops are longer than postvocalic ones, yet in Russian word boundary effects on duration were less pronounced compared to English. Davidson & Roon [2] observed for Russian that C1 duration for stops remains constant under changes in syllable structure (C1#C2, #C1C2, #C1əC2); they conjecture that this is due to the obligatory release. Our own recent work on cluster timing [3] has confirmed that Russian clusters have a greater lag between the consonants than typically observed for English or German, yet surprisingly there is nonetheless a high degree of consonant-consonant coarticulation in that there is a comparatively early movement onset of C2 during C1 in a #C1C2V cluster, possibly pointing at a slower constriction formation of C2 relative to C1 (an aspect not directly analyzed in that study). It has also been reported previously based on acoustic measurements that internal cluster members (C2) are shorter than the external ones (C1) [17, 18]. Yet the articulatory work of Marin et al. [3] suggests that the duration of the internal cluster member (C2) may actually be longer than the duration of C1 if closure formation is taken into consideration (which can only limitedly be done acoustically).

In the present paper, we examine the durational characteristics of Russian onset clusters in terms of both constriction formation and plateau duration of the two consonants involved. We further determine the relative timing patterns of C1, C2 in terms of both plateau lags as well as in terms of the onset of constriction formation for C2 in relation to C1. On the basis of these measures, we ask whether speaking rate changes are implemented similarly to what has been reported for American English clusters, i.e. as a combination of consonant-consonant timing and durational changes of each individual consonant. If the low overlap pattern is implicated in phonotactics as argued by Pouplier & Beňuš [8], we predict that speaking rate effects in Russian onset clusters are largely confined to durational changes of the individual consonants rather than affecting changes in C1-C2 timing (plateau lag). This would predict that rate effects are most clearly observable in terms of when the constriction formation of C2 begins relative to C1. Secondly we investigate to which extent the sonority profile of a cluster conditions different overlap patterns, which in turn may imply cluster-specific flexibility under speaking rate.

2. Method

We report articulography (EMA, Carstens AG501) data for five native speakers of Russian. The kinematic data were sampled at 12kHz, synchronized audio at 25.6kHz. Standard calibration and position recovery methods were used. Receiver coils were attached to the tongue (tip, mid, back), jaw, upper and lower lips, plus four reference sensors allowing for the correction of head movement. The kinematic data of the tongue tip coil were filtered at 40Hz, reference coils at 5Hz, and all others at 20Hz. For each speaker we rotated the data to the occlusal plane and obtained an outline of the hard palate.

2.1. Stimuli and Rate Elicitation

Twelve #C1C2V onset clusters were recorded with V=/a/ except for V=/o/ for /mn-, kt-/ due to phonotactic constraints. All stimuli were possible syllables of Russian. Clusters either showed a rising, flat, or falling sonority profile (see Table 1). The target syllables were embedded in the carrier phrase: [ˈgromka ____ paftaˈril], '(He) repeated ___ loudly'. Two speaking rates were elicited: Speakers saw a horizontal bar moving across the screen that indicated the time window within which they had to say the entire utterance. Bar duration was 2.3 and 1.5 seconds for the slow and fast rate respectively.

Table 1. *Stimuli by Profile Group*

Rising	Flat	Falling
bla	tka	lba
gla	kto	lga
xma	mno	mxa
ʃma	mla	ʃpa

Five repetitions per rate were recorded; the rates were blocked and speakers were alerted to rate changes. The first stimulus of each rate block was a dummy and served to remind speakers of the targeted rate. Each repetition of the two rate blocks was preceded by a block of stimuli for a different experiment; whether the slow or fast condition came first was varied for each repetition. The targeted token total amounted to 5 speakers x 12 clusters x 5 repetitions x 2 rates = 600. Data loss occurred for 20 tokens, leaving 580 tokens for analysis. 5 of these 20 tokens were individual repetitions; 15 of the 20 tokens stem from S4, for whom the /mx-/ cluster is missing completely for both rates, and /xm-/ for the fast rate.

2.2. Measurements and Statistics

For each tongue sensor, constriction degree was computed as minimal Euclidean distance to the palate; for the lips as Euclidean distance of upper to lower lip. Constriction formation and target achievement were determined for each consonant by a 20% threshold of the peak velocity of the constriction formation; the release time point was computed relative to the peak velocity of the constriction release.

We assessed speaking rate effects in terms of the duration of the consonants (closure formation and plateau) as well as the relative timing of the two consonants to each other. For the latter we employed two measures: one measuring consonant coarticulation, and one measuring plateau lag between C1 and C2. The particular *coarticulatory overlap* measure used here captures when, during the articulatory target plateau of C1, constriction formation of C2 begins. We calculate the time lag between target achievement of C1 and begin of constriction formation for C2 relative to the plateau duration of C1. This measure expresses which percentage of the C1 target plateau occurs before movement onset of C2, therefore smaller values mean more coarticulatory overlap. Negative values mean that the movement of C2 is initiated before C1 reaches its target. *Plateau lag*, our second measure, was computed as the time point of target achievement of C2 minus release of C1. *Plateau duration* of each consonant was defined as time point of release minus time point of target achievement of a given consonant. *Constriction formation* was defined as the time interval between onset of movement and target achievement of a given consonant.

For statistics we used linear mixed models. P-values were obtained by a likelihood-ratio-test in which the full model was compared to a model without the factor in question.

3. Results

Global Rate Implementation. The actual speaking rate for each trial was determined based on a syllables/second calculation. We measured for each trial from the acoustic signal the duration of the entire sentence and calculated speaking rate as syllables per second. The across-subject syllables per second mean was 5.44 (SD = 0.53) for the fast and 4.6 (SD = 0.64) for the slow condition. A mixed linear model on the dependent variable syllables per second (fixed factors: Rate, Cluster, random: Speaker, Repetition) rendered $X^2(1) = 667.7$ $p < .001$. The interaction Rate*Cluster was not significant $X^2(11) = 11.04$ $p = .44$. Thus subjects succeeded in contrasting two speaking rate conditions.

3.1. Interval Durations: Durational Characteristics of C1 and C2 as a Function of Speech Rate

For each cluster member C1, C2 we look at the duration of constriction formation (CLO) and at plateau duration (PLAT). Table 2 gives the averages in ms for both measures for C1 and

C2. Two mixed models were run, one for each measure as dependent variable. Independent variables were Consonant (C1, C2) and Rate (slow, fast), with random slope and intercept for Speaker and Cluster and random intercept for Repetition. For both dependent measures, both main effects were significant but not the interaction (Closure: Consonant $X^2(1) = 115.92$, $p <.001$; Rate $X^2(1) = 32.41$ $p < .001$; Interaction: $F = 1.03$; Plateau: Consonant $X^2(1) = 54.4$ $p <.001$; Rate $X^2(1) = 106.25$ $p < .001$; Interaction: $F = 0.99$).

Table 2. *Constriction formation (CLO) duration and plateau duration (PLAT) in ms for C1 and C2 by rate.*

RATE	C1 CLO	C2 CLO	C1 PLAT	C2 PLAT
SLOW	80.08	93.00	55.22	43.47
FAST	75.08	85.79	39.58	30.48
TOTAL	77.59	89.38	47.43	37.00

The previously reported durational contrast for internal and external cluster members in our data holds for the plateau duration measure only: The internal cluster member (C2) is shorter than the external member. For closure formation (CLO), the relationship is, however, the opposite: C2 has an on average longer constriction formation than C1. The lack of an interaction confirms that these durational patterns are stable across speaking rates. Across rates, the ratio of constriction formation duration to plateau duration is 1.6 for C1, but 2.5 for C2. These global results suggest that C1 has a relatively shorter constriction formation phase but a longer plateau, while C2 has a relatively longer constriction formation phase but a shorter plateau. At first blush, this may be attributed to intrinsic consonant duration differences, but recall that our corpus contains obstruents as well as sonorants in either position. The difference in duration ratios also means that across all the data, closure formation duration does not predict plateau duration equally well for C1 and C2. Figure 1 presents a scatter plot for C1 and C2 each, across all clusters and subjects, plotting closure formation duration against plateau duration. Interestingly, there is a positive correlation for C1 for both rates, but little correlation for C2 at either rate. At the fast condition the pairwise Pearson correlation coefficient for C1 r=0.34, for C2 r=-0.01; for the slow condition for C1 r= 0.25 and for C2 r= -0.04.

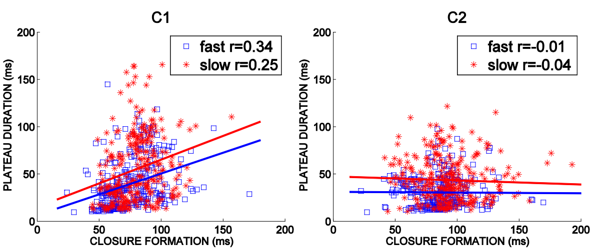

Figure 1: *Scatterplot of closure formation against plateau duration for C1 and C2 coded by rate.*

The presence of a correlation between closure formation and plateau duration for C1, but the lack of such a correlation for C2 suggests possible trading relations between C1 and C2. Conceivably, durational variation within C1 may be compensated for in the closure formation of C2 in order to ensure a certain plateau lag between the two clusters. To test for this possibility, we compute as a next step the total duration of C1 as the sum of constriction formation and plateau duration, and correlate that duration with closure formation duration of C2. The expectation is that C2 closure formation would vary with C1 total duration. Since the correlation coefficients for the two rates were very similar for C1 and C2, we collapse the data across the two speaking rates here. The pairwise Pearson correlation coefficient for C1 total duration and C2 closure formation was $r = -0.06$ ($p = .14$). Trading relations between C1 and C2 could thus not be confirmed, the constriction formation duration of C2 does not co-vary with the total duration of C1. The timing between the two consonants will be considered in detail in the next Section.

3.2. Consonant Timing: Coarticulation and Plateau Lag by Rate and Profile Group

In this section we consider how the timing between the two cluster members differs as a function of their sonority profile, as a function of speaking rate and a possible interaction between these two factors. As laid out in the Introduction, Russian stops are obligatorily released and it has been proposed that the existence of typologically rare cluster phonotactics implicates a large plateau lag between clusters. Both of these factors lead us to hypothesize that the C1-C2 plateau lag should not vary as a function of speaking rate. Figure 2 gives the C1-C2 plateau lag results by profile group and rate, Figure 3 for coarticulatory overlap.

The plateau lag results in Figure 2 reveal that against our predictions, plateau lag changes with rate. The clusters with a falling sonority profile show the smallest plateau lag and the least change with rate. Overlap values in Figure 3 are predominantly negative, meaning that C2 begins its movement into the constriction before C1 has reached its target. We ran a mixed linear model for each dependent variable with the fixed factors Rate (slow, fast) and Profile (levels: rising, flat, falling) and random factors Speaker, Cluster and Repetition. The results for plateau lag rendered a marginally significant effect for Profile group ($X^2(2) = 4.97$, $p = .08$), but a clearly significant Rate effect ($X^2(1) = 22.91$, $p<.001$). The Interaction was not significant ($X^2(2) = 4.57$, $p = 0.1$). For coarticulatory overlap (Figure 3), Profile ($X^2(2) = 3.53$, $p = 0.17$) was not significant either, but Rate was ($X^2(1) = 64.1$, $p < 0.01$), as was the Interaction ($X^2(2) = 13.24$, $p < 0.01$). A Tukey posthoc confirmed that for the rising and flat sonority profile groups there is a significant difference in overlap between the speech rates, but not for the falling sonority profile group. In sum, our results show that plateau lag significantly decreases with rate and coarticulatory overlap significantly increases with rate (recall that smaller values mean more overlap). Speaking rate interacts with profile group in that the falling sonority profile group does not display variation in coarticulatory overlap as a function of rate.

In order to test whether the rate conditions were implemented locally for all profile groups, we ran a linear mixed model on vowel duration as dependent variable and Profile Group and Rate as independent variables (random factors Speaker, Cluster, Repetition). Vowel duration was defined operationally as duration of the jaw opening plateau. The Rate effect was significant ($X^2(1) = 64.09$, $p < .001$), but crucially not the interaction ($X^2(2) = 0$). That is, all three profile groups showed a similar implementation of speaking rate in the vowel.

We end our presentation of results with example data from each sonority profile group. Figure 4 gives as an example average normalized bar graphs for the time course of C1, C2 closure formation and plateau duration for the clusters /bl-, lb-, tk-, kt-/. These example data illustrate how C2 constriction formation during /bl-/ is much earlier in C1 compared to /lb-/, but also how plateau lag is greater in /bl-/ than /lb-/. The plateau lag in /bl-/ changes as a function of speech rate more than the plateau lag in /lb-/ does (patterns also observed in the overall data). A similar difference between /tk-/ and /kt-/ (both flat sonority profile clusters) underscores that the /bl- vs lb-/ variation is not (exclusively) due to sonority profile. /kt-/ is more similar to /bl-/ in its C1-C2 timing, but the reverse order /tk-/ is more similar to /lb-/. These examples suggests that segmental effects beyond sonority (e.g. possibly an effect of C2 being a coronal) may play a more important role in determining cluster timing than sonority.

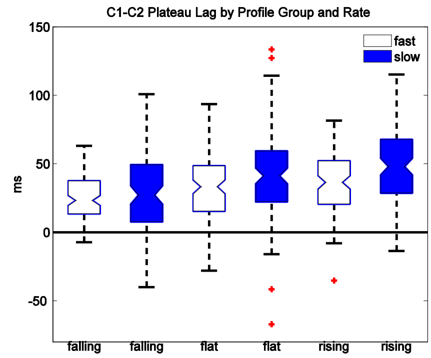

Figure 2: *C1-C2 plateau lag by Profile Group and Rate.*

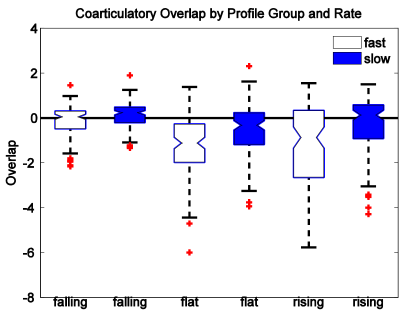

Figure 3: *Coarticulatory overlap by Profile Group and Rate. Smaller values mean more overlap.*

4. Discussion

The goal of our study was to investigate the durational characteristics of Russian consonant clusters, how they vary with speaking rate and whether rate interacts with the sonority profile of the cluster. Our study was motivated by previous research which had revealed that Russian onset clusters show a low degree of plateau overlap and only limitedly vary in duration as a function of prosodic affiliation, as would be expected from an English perspective. Relatedly, low overlap patterns have been hypothesized to be a conditioning factor in typologically unusual consonant phonotactics, opening the possibility that also within Russian, the sonority flat and falling clusters might be less susceptible to changing their timing characteristics with rate.

Our interval duration results indicated that the known durational asymmetry within a cluster [17, 18] with the internal member being shorter than the external one only holds for plateau duration. For constriction formation duration C2 has in fact a longer formation duration than C1. There was also some degree of co-variation of constriction formation and plateau duration for C1, but not for C2. We could, however, not trace this to a 'compensatory' effect of C2 constriction formation varying with C1 duration. The independence of plateau and constriction formation duration for C2 is surprising from a motor control perspective: if duration is, as often assumed [12], the consequence of a consonant's intrinsic stiffness parameter, closure formation and plateau should scale uniformly. This was, however, not the case for C2. Models advocating a separate closure and release gesture for consonants [19-21] would not predict a correlation between constriction formation and plateau duration. Nonetheless also for these models it remains for now unresolved why we would see a co-variation for C1 but not C2. While our results suggest interactional effects between the two consonants, these effects could not be pinned down in the current study.

The hypothesis that speaking rate changes would not affect C1-C2 plateau lag could not be confirmed. A faster speaking rate globally conditioned a shorter plateau lag and more coarticulatory overlap. As to the sonority profile of a cluster, an interaction with speaking rate became evident in that the falling sonority profile group (/lb-, lg-, ʃp-, mx-/) showed no change in overlap with rate. The falling sonority profile group was also the one with the overall smallest plateau lag (although statistically this was only marginally significant). This may suggest that these clusters are the least flexible in their timing because they show overall less coarticulation than the other cluster groups. However, our examination of specific examples in which sonority-flat /kt-/ patterns with sonority-raising /bl-/, but sonority-flat /tk-/ with falling /lb-/ underscores the fact that segmental composition effects beyond sonority may play an important role in shaping cluster timing. Such effects remain to be explored in future work.

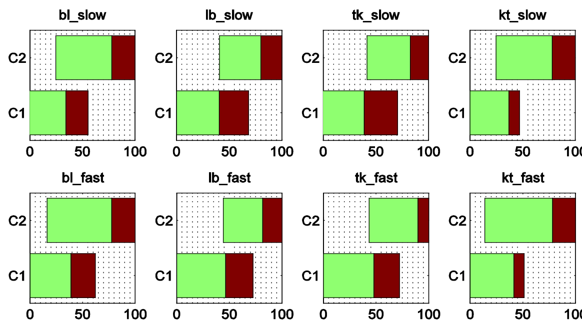

Figure 4: *C1 C2 timing. The x-axis shows normalized time, the (light) green boxes show the constriction formation interval, the (dark) brown boxes the plateau interval.*

5. Acknowledgements

We thank Susanne Waltl and Anna Ratzinger for help with data recording and segmentation. Work supported by the ERC under the EU's 7th Framework Programme (FP/2007-2013)/Grant Agreement n. 283349-SCSPL.

6. References

[1] Zsiga, E., Articulatory timing in a second language: Evidence from Russian and English. *Studies in Second Language Acquisition*, 2003. **25**(3): p. 399-432.

[2] Davidson, L. and K. Roon, Durational correlates for differentiating consonant sequences in Russian. *Journal of the International Phonetic Association*, 2008. **28**(02): p. 137-165.

[3] Marin, S., M. Pouplier, and A. Kochetov, Timing patterns of word-initial obstruent-sonorant clusters in Russian. *Proceedings of the 18th International Congress of Phonetic Sciences, Glasgow*, 2015.

[4] Bombien, L. and P. Hoole, Articulatory overlap as a function of voicing in French and German consonant clusters. *Journal of the Acoustical Society of America*, 2013. **134**(1): p. 539-550.

[5] Zsiga, E., Phonetic alignment constraints: Consonant overlap and palatalization in English and Russian. *Journal of Phonetics*, 2000. **28**: p. 69-102.

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

[7] Byrd, D. and C. Tan, Saying consonant clusters quickly. *Journal of Phonetics*, 1996. **24**(2): p. 263-282.

[8] Pouplier, M. and Š. Beňuš, On the phonetic status of syllabic consonants: Evidence from Slovak. *Journal of Laboratory Phonology*, 2011. **2**(2): p. 243-273.

[9] Adams, S.G., G. Weismer, and R.D. Kent, Speaking rate and speech movement velocity profiles. *Journal of Speech and Hearing Research*, 1993. **36**(February): p. 41-54.

[10] Lindblom, B., Spectrographic study of vowel reduction. *Journal of the Acoustical Society of America*, 1963. **35**(11): p. 1773-1781.

[11] Gay, T., Mechanisms in the Control of Speech Rate. *Phonetica*, 1981. **38**: p. 148-158.

[12] Ostry, D. and K.G. Munhall, Control of rate and duration of speech movements. *Journal of the Acoustical Society of America*, 1985. **77**: p. 640-648.

[13] Shaiman, S., E.L. Saltzman, and B. Tuller, Timing relationships of the upper lip and jaw across changes in speaking rate. *Journal of Phonetics*, 1995. **23**: p. 119-128.

[14] Smith, A., et al., Spatiotemporal stability and patterning of speech movement sequences. *Experimental Brain Research*, 1995. **104**: p. 493-501.

[15] Nittrouer, S., Phase relations of jaw and tongue tip movements in the production of VCV utterances. *Journal of the Acoustical Society of America*, 1991. **90**: p. 1806-1815.

[16] Kochetov, A., M. Pouplier, and M. Son, Cross-language differences in overlap and assimilation patterns in Korean and Russian. *Proceedings of the XVI International Congress of Phonetic Sciences, Saarbrücken*, 2007: p. 1361-1364.

[17] Redford, M., Word-internal versus word-peripheral consonantal duration patterns in three languages. *Journal of the Acoustical Society of America*, 2007. **121**: p. 1665 - 1678.

[18] Umeda, N., Consonant duration in American English. *Journal of the Acoustical Society of America*, 1977. **61**: p. 846 - 858.

[19] Browman, C., *Lip aperture and consonant releases*, in *Phonological Structure and Phonetic Form. Papers in Laboratory Phonology III*, P. Keating, Editor. 1994, Cambridge University Press: Cambridge. p. 331-353.

[20] Nam, H., Articulatory modeling of consonant release gesture. *Proceedings of the 16th International Congress of the Phonetic Sciences, Saarbrücken, Germany*, 2007: p. 625-628.

[21] Harrington, J., J. Fletcher, and C. Roberts, Coarticulation and the accented/unaccented distinction: evidence from jaw movement data. *Journal of Phonetics*, 1995. **23**: p. 305-322.