Vowel Compensatory Shortening in Romanian

Stefania Marin[1], Lia Saki Bučar Shigemori[2]

[1,2] *Institute of Phonetics and Speech Processing, Ludwig-Maximilian University Munich, Germany*

marin@phonetik.uni-muenchen.de, lia@phonetik.uni-muenchen.de

Abstract

This study systematically investigates the change in relative vowel and consonant duration as a function of increased complexity of onsets and codas (vowel/consonant compensatory shortening), and its interaction with type of onset/coda in Romanian. Results show that vowels shorten when they follow complex stop and liquid onsets, but not nasal or fricative ones, and when they precede liquid complex codas, but not nasal or obstruent ones. Compared to their timing as singletons, only liquid (but not nasal or obstruent) consonants shorten in both onset and coda position. This pattern contradicts the general predictions of either a compression or gestural account.

Keywords: compression effects, compensatory shortening, syllable temporal organization, clusters

1. Introduction

Vowel compensatory shortening (i.e. vowel shortening due to increasing syllable complexity, henceforth VCS) has been reported in several studies (Katz 2012, Marin & Pouplier 2010, Munhall et al. 1992, Shaiman 2001). The conditioning environment, as well as the theoretical interpretation of VCS, have however been a matter of debate. On the one hand, VCS has been viewed as one instance of compression effects, whereby segments in syllables of greater complexity tend to shorten relative to their duration in simpler syllables, so that the overall syllable duration remains fairly regular (cf. Munhall et al. 1992). Under this hypothesis, vowels are expected to be shorter in syllables with complex onsets or codas, compared to vowels with simple margins.

On the other hand, under a gestural approach to syllable organization (Browman & Goldstein 1998), VCS has been hypothesized to be a consequence of a temporal organization specific to onsets, but not to codas, and hence VCS is predicted when onset but not when coda complexity increases. Specifically, onsets are hypothesized to be timed globally to the following vowel, i.e. the onset's mid-point (so called "c-center") is assumed to maintain a stable relationship to the vowel. Consequently, when consonants are added to the onset, the rightmost and leftmost consonants are assumed to shift rightwards and leftwards respectively, relative to their timing as a singleton, so that onsets of varying complexity line up along their c-centers. This shift would result in an increasing overlap of the rightmost consonant and the following vowel, resulting in an apparent vowel shortening in the complex vs. simplex onset condition (i.e. an apparent VCS). No such increasing overlap is expected for codas, since they are hypothesized to be timed locally, i.e. only the vowel-adjacent consonant is assumed to be coordinated to the preceding vowel, and this timing should not be affected by increasing coda complexity.

The empirical evidence on VCS is partly contradictory even for a single language, English. For onsets, VCS has been reported independently of onset composition (Katz 2012, Marin & Pouplier, 2010), a pattern compatible with the predictions made by either a compression or a gestural account. The evidence for coda VCS is however mixed across studies and coda types. Thus, VCS in the context of obstruent codas such as /-ps/, /-sp/ has been observed by Munhal et al. (1992) and Shaiman (2001), but not by Byrd (1995). More recent studies (Katz 2012, Marin & Pouplier 2010) have corroborated the result of Byrd (1995) for obstruent codas, while at the same time reporting VCS for liquid codas (such as /-lp/, /-rp/), a pattern not predicted by either a compression account (which would predict VCS independently of coda type) or a gestural approach (which would predict no VCS as a function of coda complexity increase). Under either approach, additional mechanisms would have to be postulated to account for the empirical English coda pattern (see e.g. Katz 2012). In the present study, we test to what extent the English VCS results are extensible to a new language, Romanian, and we specifically investigate the role of onset/coda composition on VCS.

Additionally, we also investigate consonant duration as a function of syllable complexity. Under a compression approach, consonant compensatory shortening (CCS) is also expected to be a mechanism available in maintaining comparable durations between simpler and more complex syllables (in addition to VCS, or in a trade-off relation with VCS). Previous research on English has suggested an overall tendency for consonants in clusters to be shorter than their singleton counterparts (Haggard 1973), although the effect was dependent on both consonant type involved and position in the syllable and cluster: thus, for example /l/ shortened in both onset and coda, but /m/ only shortened in coda if it was the first consonant in the cluster (e.g. /m/ shortened in coda /-md/, but not in coda /-lm/, or in onsets /sm-/); likewise stops such as /p/ shortened if they were the second consonant in a coda cluster (/-lp/), but not the first (/-ps/). A different study suggested an overall shortening of consonants in onset clusters, but a lengthening of consonants in coda clusters (O'Shaughnessy, 1974), but with different degrees of shortening in onset as a function of consonant type (minimal for stops, greater for liquids). Note that in this study no statistical analyses were carried out as only three speakers were recorded in the onset condition and two speakers in the coda condition. The current data allow us to investigate the relationship between VCS and CCS and their interplay with onset/coda composition.

2. Methods

Articulatory (EMA) data from five speakers were used to compare vowel duration in cluster words with vowel duration in corresponding singletons (e.g. *gl̲as* vs. *l̲as* forming Set GL-, with /l/ being henceforth referred to as the vowel-adjacent consonant). Syllable complexity was increased in either onset (e.g. Set GL-), or coda (e.g. Set –LG: *mul* vs. *mulg*, with /l/ being the vowel-adjacent consonant in this set). Sets were further grouped on the basis of the vowel-adjacent consonant into four series: liquid (/l/, /r/), nasal (/m/, /n/ in onset, /m/ in coda position), fricative (/s/) and stop (/p/, /k/, /t/ in onset, /p/ and /k/ in coda). All available sets are listed in Table 1. All clusters were mono-morphemic and the stimuli were real words, embedded in carrier phrases. Six repetitions were targeted for each stimulus word (cf. Marin 2013, Marin & Pouplier 2014 for methodological details).

Table 1: *Experimental sets.*

Series	Onset	Coda
Liquid	PL, BL, KL, GL, FL, ML, PR, BR, KR, GR, MR	LB, LK, LG, LF, LM, RP, RB, RK, RG
Nasal	SM, SHM, KN	MN
Fricative	PS, KS	SK, SM
Stop	SP, SHP, SK, SHK, KT	PS, PT, KS, KT

Vowel (and consonant) acoustic segmentation is known to be problematic for certain segments and in certain contexts (e.g. vowels in the context of liquids, and especially the alveolar trill) and often different criteria have to be adopted on a case-by-case basis (cf. Haggard 1973, Katz 2012). In the current study, the availability of articulatory data allowed us to circumvent this problem by defining vowel and consonant durations on the basis of consistent kinematic events across contexts. Thus, using the Matlab-based *Mview* software algorithm developed by Mark Tiede at Haskins Laboratories, kinematic events defining target achievement and release of consonants were determined on the basis of changes in the velocity profiles of relevant articulator movements (cf. Marin 2013, Marin & Pouplier 2014 for details). The semi-automatic labeling procedure consisted in first detecting in the region of interest two velocity peaks corresponding to the articulator moving towards and away from target (Peak 1, Peak 2 in Figure 1). Target achievement (Target) was defined as the point in time at which velocity fell below 20% of Peak 1, and target release (Release) was defined as the point at which velocity exceeded 20% of Peak 2. All landmarks were determined automatically after manually selecting a point of analysis.

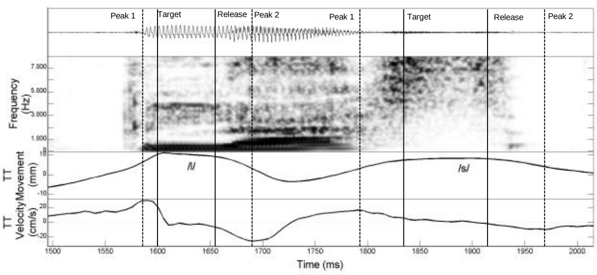

Figure 1: *Example articulatory segmentation showing vertical movement and velocity trajectories for the tongue tip (TT), and kinematic events for consonants /l/ and /s/ in word "glas". Vowel duration was defined as the interval between release of /l/ and target of /s/. Vowel-adjacent consonant duration (in this instance duration of consonant /l/) was defined as the interval between target and release of /l/.*

Vowel duration was approximated as the interval between the articulatory release of the consonant preceding the vowel and the target achievement of the consonant following the vowel (Figure 1). Consonant duration was defined as the interval between target and release of the consonant of interest (Figure 1). For evaluating CCS, only the vowel-adjacent consonant duration was compared across singletons and clusters since only this was available in both conditions.

For statistical analyses, mixed linear models were computed using the *lme4* package for R. To determine *p*-values for the main effect and the interactions between factors, a model including the fixed factor/interaction of interest was compared to the same model with no fixed factor/no interaction (cf. Bates 2010). This method circumvents the difficulty in estimating denominator degrees of freedom for mixed linear models. The *p*-values thus obtained is reported along with the F-value of the mixed linear model; because denominator degrees of freedom are difficult to estimate, they cannot be reported (moreover, they no longer play a role in computing *p*-values, given the method used here). For post hoc comparisons, the *p*-values were determined using the single-step adjusted method in the *multcomp* package for R (Hothorn, Bretz, & Westfall 2008). The data were analyzed with fixed factors Complexity (singleton, cluster), Position (onset, coda) and Series (liquid, nasal, fricative, stop), and Speaker and Set were included as random factors. Fixed factors Position and Series were of interest to the extent that they interacted with Complexity, hence we focus on reporting the results for Complexity as a main effect and its interaction with Position and Series.

The prediction under a compression account is that VCS and CCS should be observed regardless of where (in onset or coda) syllable complexity increases, and regardless of composition. Possibly, trade-off relations may be at play between VCS and CCS. A gestural approach would predict vowel shortening as a function of onset but not coda increased complexity (with no clear predictions being made for consonant duration beyond the general prediction that compensatory shortening is the result of increasing overlap rather than gestural shortening and that it should be expected to the extent that consonants in clusters overlap each other more). Taking into account the English results, the predictions on vowel duration are more nuanced, with vowel shortening being expected for all onsets regardless of their segmental composition, but only for the liquid coda sets. The predictions on consonant duration are that shortening should be observed mainly in onset condition, and more robustly for liquids than other consonant types.

3. Results

3.1. Vowel duration

Vowel durations in the singleton and cluster conditions as a function of series are plotted in Figure 2. A significant effect for factor Complexity ($F = 57$, $p<.001$), for its interaction with Position ($F = 8.24$, $p=.004$) and Series ($F = 9.6$, $p<.001$), as well as for the three-way interaction ($F = 5.43$, $p<.001$) was observed. Post-hoc tests showed that in onset position, liquid and stop sets had significantly shorter vowel durations in the cluster compared to the singleton condition ($p<.007$). No significant difference in vowel length between conditions was observed for the nasal onsets ($p>.05$), and a significant vowel lengthening in the cluster condition was observed for the fricative onsets ($p=.015$). In coda, the vowel was shorter in the cluster condition only for the liquid sets ($p<.001$), with no significant difference being observed for the other series ($p>.05$; the comparison for the fricative series was at trend level, $p=.057$). The results therefore confirmed VCS for liquid and stop onset sets, and for liquid (and marginally fricative) coda sets.

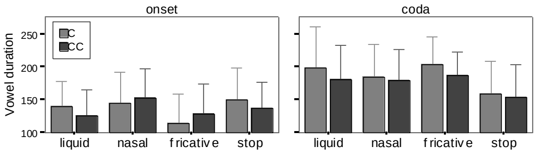

Figure 2: *Mean (+1SD) vowel duration (ms) in the cluster (CC) and singleton (C) condition, per position and series.*

3.2. Consonant duration

Vowel-adjacent consonant durations in the singleton and cluster conditions as a function of series are plotted in Figure 3. For consonant duration, factor Complexity (F = 29, *p*<.001), and the two-way interactions with Position (F = 5.48, *p*=.019) and Series (F = 2.76, *p*=.041) were significant, but not the three-way interaction (F = 0.35, *p*>.05). Post-hoc tests exploring the two-way significant interactions showed that across all series, the consonants were overall shorter in complex onsets than in simple onsets (*p*<.001), with no complexity effect in coda position (*p*>.05); across syllable positions, liquids in clusters were shorter than singleton liquids (*p*<.001) , with no other series effects (*p*>.05). The absence of a three-way interaction suggests that CCS for liquids should be observed independently of position, and indeed a post hoc analysis confirmed that in both onset and coda position, liquids in the cluster condition were significantly shorter than singleton liquids (Liquid onsets: *p*<.001; Liquid codas: *p*=.037).

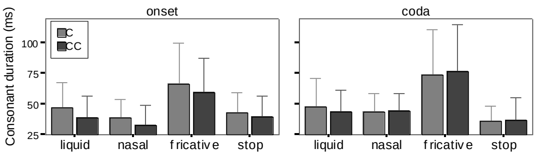

Figure 3: *Mean (+/- 1SE) consonant duration (ms) in the cluster (CC) and singleton (C) condition, per position and series.*

4. Discussion

In Romanian, both VCS and CCS were observed for selected syllable types: VCS was robustly observed for the liquid and stop onsets and for the liquid codas, while CCS was overall observed for liquids, and in onset position; possible trading relations between VCS and CCS could be observed for the nasal and fricative onsets, where the consonant but not the vowel shortened as a function of syllable complexity. Interestingly, when syllable complexity increased in coda, CCS was only observed in contexts where VCS was also present, i.e. within the liquid series. This pattern contradicts the general predictions of a compression account, and overall more conditions were found under which segmental compression was blocked than observed. This suggests that segmental compression to the extent that it is present, should be viewed as an epiphenomenon rather than as a driving mechanism in speech production.

At the same time, a gestural account to syllable organization also does not entirely predict the observed vowel duration pattern. Specifically, the vowel shortening pattern for liquid codas, while mirroring the pattern reported for English, does not follow from the gestural organization hypothesized for codas. In addition, fricative and nasal onsets do not conform to the predictions made for onset temporal organization: for the fricative sets, the vowel lengthens rather than shortens with increased syllable complexity, while no difference in vowel duration is observed for the nasal series. In addition to the principled distinctions between the organization of onsets and codas, additional factors would have to be considered to account for the unexpected patterns in the fricative and nasal onset conditions, as well as in the liquid coda condition.

Onset clusters where the vowel-adjacent consonant is a fricative (/ps-/, /ks-/) are not attested in English, so it is not clear whether the onset fricative pattern is language- or cluster-specific, although recent data from Polish (Pastätter & Pouplier 2014) suggests that it may be cluster-specific. Given the known coarticulatory resistance of the sibilant (Recasens 2012), it could be envisioned that the sibilant and the vowel cannot overlap more in the cluster condition than they already do in the singleton condition, and hence the sibilant cannot shift rightwards relative to its timing as a singleton, as it would be expected for complex onsets (cf. Pastätter & Pouplier 2014, for a similar suggestion for Polish sibilant onset clusters). What is however surprising in the Romanian result is that the vowel is significantly longer in the cluster condition, suggesting that the sibilant and vowel overlap even less in the cluster compared to the singleton condition. At this point it is not clear why that is the case. In Romanian, these fricative clusters are extremely rare, and it may be that they are produced differently than regular onsets, or produced entirely differently than their singleton controls (cf. Marin 2013).

The nasal onset pattern is surprising since it not only diverges from the pattern predicted for onsets, but also from the pattern observed for English nasal onsets. As far as can be determined, nasals in the two languages are not different in any crucial way, so it is not straightforward to propose an analysis that would predict the two patterns in the two languages. When examining the sets in more detail, it can be seen that there is on average a vowel shortening for set SM- (which is similar to SM-, SN- in the previous English studies), but not for sets SHM- and KN-. Like in the case of /ps-/, /ks-/, clusters /shm-/ and /kn-/ are rare in Romanian, so again they may not be treated like regular onsets (indeed, in a different type of analysis, there were suggestions that sets PS-, KS-, KN- in Romanian pattern together, cf. Marin 2013).

Finally, the liquid coda pattern mirrors that of English, and for English additional perceptual factors were postulated to explain why the vowel shortened in this context (Katz 2012, Marin & Pouplier 2010). Alternatively, production particularities characterizing liquids but not nasals or obstruents may result in an apparent vowel shortening given the measure employed here (as well as acoustic vowel segmentation conventions). Thus for example, the tongue dorsum retraction needed for producing the Romanian trill, which precedes the tongue tip constriction, is part of the measured "vowel" (either in an articulatory measure as we use here, or in an acoustic measurement). If this tongue dorsum retraction gesture shortens, just like the tongue tip constriction has been shown to shorten, in the cluster vs. singleton condition, then the measured "vowel" would also automatically shorten. If so, the observed VCS for liquid codas is actually an artifact of the imperfect articulatory and acoustic measures available to us (cf. Marin & Pouplier 2014 for a more comprehensive discussion in the context of various cross-linguistic liquid clusters).

The results overall highlight the need for empirical data from a variety of languages and onset/coda types to evaluate competing theoretical assumptions regarding vowel and consonant compensatory shortening.

5. Acknowledgements

Work supported by the European Research Council under the European Union's Seventh Framework Programme/ERC Grant Agreement no. 283349-SCSPL.

6. References

Bates, D. (2010). *lme4: Mixed-effects modelling with R*. Springer.

Browman, C. P., & L. Goldstein (1988). Some Notes on Syllable Structure in Articulatory Phonology. *Phonetica*, *45*, 140–155.

Byrd, D. (1995). C-centers revisited. *Phonetica*, *52*, 285–306.

Haggard, M. (1973). Abbreviation of consonants in English pre- and postvocalic clusters. *Journal of Phonetics*, *1*(1), 9–24.

Hothorn, T., F. Bretz, & P. Westfall (2008). Simultaneous Inference in General Parametric Models. *Biometrical Journal*, *50*(3), 346–363.

Katz, J. (2012). Compression effects in English. *Journal of Phonetics*, *40*(3), 390–402.

Marin, S. (2013). The temporal organization of complex onsets and codas in Romanian: A gestural approach. *Journal of Phonetics*, *41*, 211–227.

Marin, S., & M. Pouplier (2010). Temporal organization of complex onsets and codas in American English: Testing the predictions of a gestural coupling model. *Motor Control*, *14*, 380–407.

Marin, S. & M. Pouplier (2014). Articulatory synergies in the temporal organization of liquid clusters in Romanian. *Journal of Phonetics*, 42, 24-36.

Munhall, K. G., C. A. Fowler, S. Hawkins & E. Saltzman (1992). "Compensatory shortening" in monosyllables of spoken English. *Journal of Phonetics*, *20*, 225–240.

O'Shaughnessy, D. (1974). Consonant durations in clusters. *IEEE Transactions on Acoustics Speech and Signal Processing, 22,* 282-295.

Pastätter, M. & M. Pouplier (2014). The temporal coordination of Polish onset and coda clusters containing sibilants. *Proceedings of the 10th International Seminar on Speech Production, Cologne, Germany, 5-8 May 2014.*

Recasens, D. (2012). A study of jaw coarticulatory resistance and aggressiveness for Catalan consonants and vowels, *Journal of the Acoustical Society of America, 132*, 412-420.

Shaiman, S. (2001). Kinematics of compensatory vowel shortening: the effect of speaking rate and coda composition on intra- and inter-articulatory timing. *Journal of Phonetics*, *29*, 89–107.