Seminar: *Atelier in Experimental and Computational Phonology*
Prof. Ph.D. Adamantios Gafos

The role of stress in the temporal organisation of Spanish syllable onsets

Antonia Schmidt

antonia.schmidt.2@uni-potsdam.de

Matrikelnr.: 8152706

November 7, 2022

1 Introduction

This work is concerned with the effects of stress on syllable intern timing relations in Spanish. Different studies (e.g. Gibson et al. (2019), Sotiropoulou et al. (2020)) have been looking at Spanish syllable organization under different view points and with different independent variables such as voicedness, manner of articulation of consonants and vowel. The question in this work is whether Spanish syllabic organization is influenced by stress. Here, the focus will lie on temporal stability (global vs. local syllable onset timing), although Sotiropoulou et al. (2020) have shown that this metric alone can not always be sufficient to describe syllable internal reorganization between complex and simple onsets.

By working with Electromagnetic Articulometry data (EMA data), I examined the variability of the duration of intervals between a postvocalic anchor and landmarks in the tautosyllabic consonants before this vowel depending on stress. Since stress does increase syllable duration, but theoretically longer durations are mitigated by calculating the relative standard deviations, I did not expect noticeable differences. Contrary to that, there seem to be some differences between stressed and unstressed syllables such as that stressed syllables show greater variability in timing relations and also sometimes exhibit locally timed onset clusters.

2 Related Work

The question of how evidence for the relatively abstract concept of a syllable can be found in a continuous phonetic signal is relatively old. Looking at the timing relations between the onset and the rhyme can show underlying mental representation of syllables.

Global vs. local timing Phonotactics of languages differ in how syllables can be constructed and what kinds of clusters are allowed where in the syllable. Complex syllable onsets such as /ksl/ may be a valid onset in a language but not in another, while two languages where this onset is valid still differ in their temporal organization of those consonants with respect to the nucleus vowel of the syllable. There have been

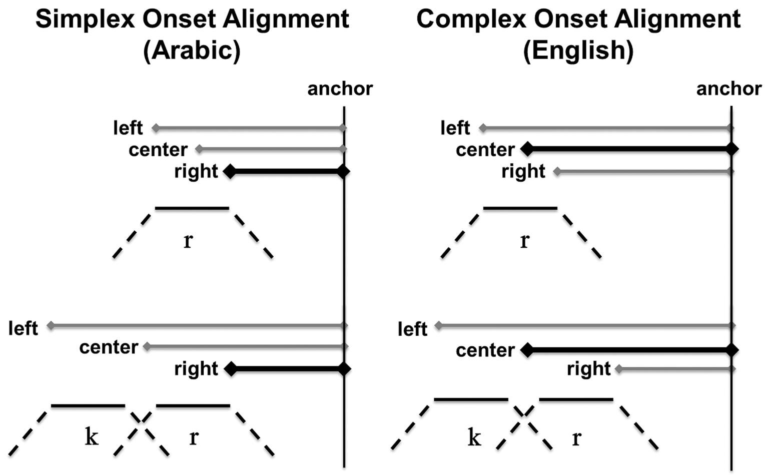

Figure 1: Local (Arabic) and global (English) consonant cluster organization: prototypes for simple and CC-cluster onsets showing the respective interval stabilities. From Shaw and Gafos (2015), CC BY 4.0.

found two different kinds of consonantal onset organization:
"Global timing refers to the coordination of a sequence of consonants as a unit with respect to the vowels on either side. Alternatively, 'local timing' refers to the coordination of a single consonant gesture in a sequence with respect to a neighboring vocalic gesture. It has been hypothesized by (Browman and Goldstein, 1988) that both local and global timing play a role in coordinating consonant sequences with adjacent vowels." (Byrd, 1995, 286)

A tool to determine which organization underlies the syllable of a language is the notion of stability. Depending on whether syllable onsets are organized locally or globally, the duration of intervals between the relevant landmarks and a (post-)vocalic anchor varies more or less when more consonants are added to an onset cluster. Landmarks found to be relevant are the right edge (i.e. the offset of the last consonant before the vowel) as for example in Arabic (Shaw and Gafos (2015)) and the c-center the midpoint between the target and the release of the syllable onset as for example in English (Shaw and Gafos (2015)). Global timing relates to c-center stability and local timing to stability of the right edge to anchor interval. This stability measure can give a good hint at how syllables in languages are organized, but it is not always sufficient. The concept is illustrated in Fig. 1 (Shaw and Gafos (2015)). Besides the stability metric (Sotiropoulou et al., 2020) finds other, more nuanced parameters which will be discussed below.

Spanish Gibson et al. (2019) looked at CCVC-syllables in Spanish, at their VOT, gestural plateau duration and gesture overlap. Concerning stability they found patterns that seemed to imply that Spanish does not allow complex syllable onsets (robust IPI) as Arabic does, which organizes consonant clusters locally. But, as opposed to Arabic, Spanish has been found to apply overall global timing by sotiropoulou2020global. In their paper Sotiropoulou et al. (2020) they found that the stability measure is not a reliable tool to detect the underlying syllable organization of Spanish. Other factors such as voicedness, and manner of articulation can play into the temporal organization of the tongue movements. For instance, the interval stabilities seemed to be quite variable; in voiced stop-lateral clusters, local timing interval stability was found and global timing interval stability in the other instances. So, only a stability metric is not enough to determine a general rule for a language, this is evidence "that syllable structure does not have consistent phonetic manifestations in the articulatory record." (Sotiropoulou et al., 2020, 22). For sake of the simplicity of this paper, I will nonetheless focus on only stability and the influence of stress while being aware of potentially contradicting results.

Stress In general, according to Prieto and Roseano (2018), Spanish has a lexical stress / word stress. This can be demonstrated in minimal pairs - or even minimal triples such as in example (1). Here, stress is represented by capitalization.

(1) TERmino - term, end
terMIno - finish.1SG.PRS
termiNO - finish.3SG.IPFV

These stressed syllables are marked acoustically by 'longer durations, higher fundamental frequency and higher intensity than unstressed syllables' (Prieto and Roseano, 2018, 215) although the overall pitch can vary depend on the position of the word in the whole phrase or sentence.
There have not been yet studies on the effect of stress in the temporary organiza-

tion and articulation of Spanish syllables, but results from other languages and the influence of stress on articulation could give us a hint. For instance, in English it has been found by De Jong et al. (1993) that coarticulation is reduced in English stressed syllables and locally the articulation is shifted towards the hyperarticulate end of the continuum. A potential pitfall is that that while they "define stress as the set of prosodic categories which involve relationships of relative prominence between syllables."(De Jong et al., 1993, 199) they look at nuclear accent on the whole phrase level in their analysis, which is not exactly what I will be doing (see Section 3). So while intergestural timing is sensitive to the position of the gesture within the prosodic structure of the utterance the influence of stress at the word level is not completely clear.

Hypotheses We know that stress probably will positively influence the timing of syllables since carefully articulated gestures are also longer in duration. I still suppose that the ratio of segment length to intergestural timing will remain the same, i.e. stress will not have an influence on the temporal syllable organization.

3 Methodology

The set of stimuli for this study is a part of a larger corpus of Spanish. The annotated and analyzed subcorpus contains stimuli with word-medial clusters. Stimuli with word initial clusters, where the first syllable is the target of investigation are also part of the whole corpus but not of this analysis. To assess the effects of syllable segments, the following variables were varied systematically:

(i) voicedness of the first consonant in clusters, which are in this case the bilabial plosives /p,b/

(ii) the vowel in the nucleus being either /a/ or /e/

(iii) whether or not the target syllables were stressed or not.

These stimuli were in pairs of CVC- and CCVC- syllables, where second consonant of the cluster (and only C of the CV - syllables) was the lateral /l/. The coda was always

		unstressed		stressed	
		CVC	CCVC	CVC	CCVC
/bl/	a	mo**lan**	ha**blan**	mo**lar**	ha**blar**
	e	po**len**	do**blen**	so**ler**	To**bler**
/pl/	a	mo**lan**	co**plan**	mo**lar**	co**plar**
	e	po**len**	so**plen**	so**ler**	Sa**plén**

Table 1: Collection of subsection of the corpus used in this analysis. The relevant syllables are in bold.

a coronal; either /n/ or /r/. In Table 1 the pairs of words containing the relevant syllables are displayed. Spanish does not allow more than two consonants in syllable onsets and one consonant in codas ((Hirschfeld, 1983)), so an approach as in e.g. Byrd (1995) where CVC, CCVC and CCCVS sequences were compared to each other is not possible. Instead, CVC / CCVC-pairings must suffice.

The target words were embedded in the sentence *"Que diga [-word-] (a)fuera de contexto"*. In ten to eleven repetitions per word this sentence was uttered by two male Spanish Speakers (one from Valladolid and on from Galicia, also speaking Gallego, but Castellano at school and with family) while the tongue movements were recoded by 3D Electromagnetic Articulometry (EMA). The EMA sensors were placed on the lower lip, upper lip, jaw, upper incisor, tongue tip, tongue mid and tongue back of the participants. The relevant time stamps were retrieved by labelling the movements of the relevant articulators (i.e. the relevant sensors) with respect to the segments of the syllables. The relevant articulators were: the tongue tip for /n/ and /t/ and the lower and upper lip for /p/ and /b/. For the labials, the lip aperture (LA, the magnitude of the space between the two lip sensors) was analyzed. The lower lip movement (LL) was taken into consideration as well, but since LA and LL almost always yielded the same timestamps, I chose to only include LA in my analysis.

The labelling of the gestures was done with the help of the mview software, which is a set of scripts developed by Mark Tiede at Haskins Laboratories, implemented in Matlab. The goal was to label the gestural on- and offset, the nucleus on- and offset (target and release) and the point of maximal constriction (midpoint). A gesture label is found by the algorithm by searching two local tangential velocity peaks around the supposed midpoint of a gesture defined within a user-specified time interval. These peaks should be artifacts of the opening and closing parts of an articulatory gesture of the articulator trajectory. The gestural onset was chosen as a point before the first velocity peak, where the velocity exceeded a certain threshold (20%) of the first peak velocity, the target as a point where the velocity fell below the threshold after the first velocity peak. The release and gestural offset were marked analogous to this at the second velocity peak and the point of maximal constriction was marked at the midpoint between target and release.

The following landmarks were measured: left edge as the target of the (first) consonant in the onset, right edge as the release of the (second) consonant in the onset. The so-called c-center is the midpoint between the left and the right edge. Finally, also an anchor (right delimiting landmark) to determine the intervals is measured. What part of the postvocalic gesture is chosen as anchor can differ. Sotiropoulou et al. (2020) test three different landmarks: the target of the coda consonant or its maximal constriction and the "spacial extremum" of the nucleus vowel. My analysis only includes the point of maximal constriction of the coda consonant.

A subset of the data posed some problems while labelling, because of prevocalic nasal reduction in on individual. Some coda-/n/ were difficult to label and only in the y-Dimension (lateral movements), a high enough velocity change was found to mark the landmarks of a gesture. Also in some other cases, where overall signals did not yield a result that was at least a bit congruent to the audio signal, the dimensionality of the signal was reduced to gain a label This should overall not be a problem to the analysis, since the focus here is on temporal relations which should theoretically stay the same no matter in which dimension.

To determine whether global or local organization plays a role, the stability of c-center to anchor interval and right edge to interval will be calculated with the help of the relative standard deviation (RSD) of these intervals. For the sake of completeness, also the left edge to anchor interval will be part of the analysis. The interval with the lowest RSD value will be the interval that remains the most stable over all instances. I will look at how stress influences these values and also how it influences the syllable length in this dataset. All analyses and graphic generation were carried out using R (R Core Team (2020)).

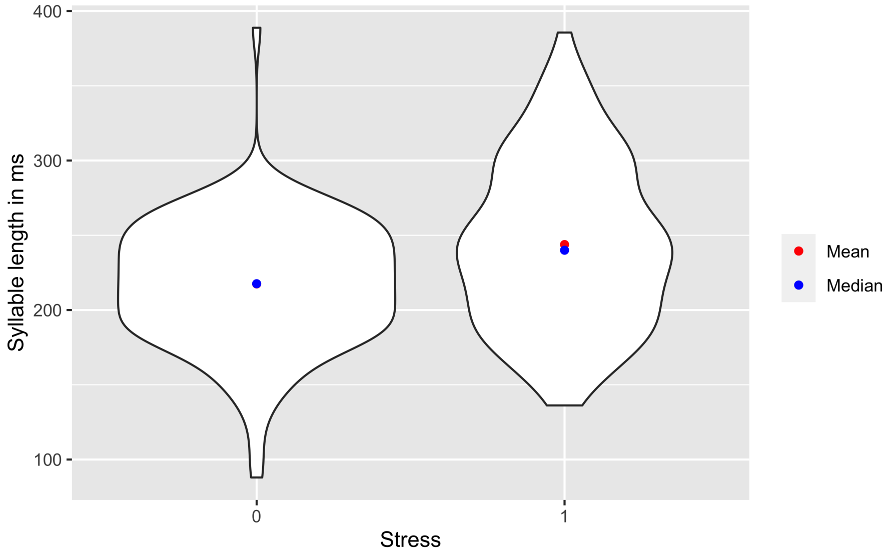

Figure 2: Violin plot of syllable length depending on stress.

4 Results and Discussion

While preprocessing, some items were excluded, either because some gestures could not be labelled or because the timing relations between the gestures deviated too much from the others. This left 256 instances of measured words.
First of all, a Welch Two Sample t-test revealed that indeed, stressed syllables are significantly ($p < 0.001$: p = 1.118636e-05) longer in duration than unstressed syllables when comparing the intervals between the left edge and the anchor. Violin plots of the distribution of syllable lengths depending on stress are illustrated in Fig. 2. The estimated mean duration value for stressed syllables is 244 ms and for unstressed syllables 217 ms. We can therefore confirm what was already stated in Section 2 after Prieto and Roseano (2018) that stressed syllables are longer than unstressed syllables and see that stress in general has an influence on syllables. In the following we will see the influence of stress on the general temporal organization of a syllable.

Table 2 shows the RSDs of the three intervals for all items, split by stress or no stress. In both cases, the c-center to anchor interval is almost always the least variable, ergo one could speak of c-center stability. However, if we look at the values broken down into Table 3, we can see exceptions to this. Interestingly, the finding of (Sotiropoulou et al., 2020) that voiced stop-lateral clusters exhibit local timing interval stability can be reproduced, but only for stressed syllables and additionally for stressed voiceless syllables, whose nucleus is /e/. Unstressed syllables show overall global timing interval stability.

However, it must also be noted that the differences between the RSD values are sometimes very small. In any case, we can observe that there is a difference between stressed and unstressed syllables, because stressed syllables generally seem to have greater variability in their interval lengths.

	Stressed_syllable	LE	CC	RE
1	yes	22.61	21.28	22.22
2	no	19.24	18.86	21.71

Table 2: Relative standard deviations of interval length between left edge (LE), C-Center (CC) and right edge (RE) and the anchor, respectively.

	Landmark	blA	blE	plA	plE	bla	ble	pla	ple
1	LE	19.66	26.04	18.64	26.64	18.76	17.97	22.13	19.82
2	CC	17.70	23.96	17.76	22.94	17.91	17.64	21.48	17.69
3	RE	17.40	22.58	21.69	21.18	18.80	20.70	23.55	20.30

Table 3: Relative standard deviations of interval length between left edge (LE), C-Center (CC) and right edge (RE) and the anchor, respectively. Broken down by syllable type; capitalization of letters indicates stress.

It could be that there is an interaction of stress and vowel, as stressed syllables with /e/ in the nucleus have the largest RSDs and unstressed voiced syllables with /a/ in the nucleus have the smallest variables, thus are show the most stable timing relations.

The fact that variability in stressed syllables is slightly higher, is surprising since I hypothesized that longer syllable durations will be compensated through normalizing with the mean when calculating the RSD and no particular differences between stressed and unstressed syllables will be found. This higher variability in the temporal organization could be because stressed syllables are more marked than their unstressed counterparts. A hypothetical explanation could be that unstressed syllables are more prototypical and 'unstressing' mitigates the influences of parameters such as voicing on the relations between gestures.

Nonetheless it has to be noted that I did not compute whether the differences between stressed and unstressed syllables are statistically significant, doing that would have been out of scope of this paper. In comparison to other papers, the calculated RSDs are relatively large. Whether this is because of the fact, that the dataset is comparatively small (only two speakers, two different vowels, only stop-lateral clusters) or because word medial clusters are different to word initial clusters, is not clear. Labelling and analyzing the rest of the corpus will surely shine some light into that.

5 Conclusion

In this paper we saw that stressed and unstressed syllables in Spanish slightly differ in their temporal organization. Stressed syllables are significantly longer than unstressed syllables, and their temporal organization shows more variation than in unstressed syllables. In contrast to unstressed syllables, stressed syllables do not uniformly show the same expected behavior of globally organized syllable onset clusters.

Whether these differences will hold when tested about statistical significance is a task for the future when more data can be included into the analysis. An analysis of the interaction of nucleus vowel and stress on the RSDs of could be promising. Furthermore more research can be done with respect to the effect of stress on the temporal relations on the timing between the consonants in CC-clusters, i.e. the interplateau interval (IPI).

References

Browman, C. P. and Goldstein, L. (1988). Some notes on syllable structure in articulatory phonology. *Phonetica*, 45(2-4):140–155.

Byrd, D. (1995). C-centers revisited. *Phonetica*, 52(4):285–306.

De Jong, K., Beckman, M. E., and Edwards, J. (1993). The interplay between prosodic structure and coarticulation. *Language and speech*, 36(2-3):197–212.

Gibson, M., Sotiropoulou, S., Tobin, S., and Gafos, A. I. (2019). Temporal aspects of word initial single consonants and consonants in clusters in spanish. *Phonetica*, 76(6):448–478.

Hirschfeld, U. (1983). Ergebnisse des Sprachvergleichs Spanisch-Deutsch im Bereich der Phonologie und Phonetik. *Deutsch als Fremdsprache*, 20(3):169–174.

Prieto, P. and Roseano, P. (2018). Prosody: Stress, rhythm, and intonation. *The cambridge handbook of Spanish linguistics*, 1:211–236.

R Core Team (2020). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.

Shaw, J. A. and Gafos, A. I. (2015). Stochastic time models of syllable structure. *PloS one*, 10(5):e0124714.

Sotiropoulou, S., Gibson, M., and Gafos, A. (2020). Global organization in Spanish onsets. *Journal of Phonetics*, 82:100995.