**Lexically specific accumulation in memory of word and segment speech rates**

## 1. Introduction

Variable pronunciations of words garner a great deal of academic scrutiny. Linguists describe and explain diachronic trajectories and synchronic patterns of variation from within a wide variety of frameworks. The abundance of research on variable articulations of words has led to a deeper understanding of the linguistic and extralinguistic factors that operate at the time of speaking to shape pronunciation variants of words.

Variationists, for example, argue that variation evident in word forms is patterned and probabilistic (Tagliamonte 2012). Numerous factors of the target context operate at the time of speaking to shape variable articulations of words. For instance, it is relatively uncontroversial that aspects of the phonetic context, such as sounds adjacent to a target, bring about well-documented adjustments to articulatory gestures resulting in patterns of greater or lesser cohesion, such as co-articulation, assimilation, lenition, and fortition. Studies of word-final t/d production in English, for example, demonstrate that rates of deletion of the coronal stop are highest when the following segment initiating the next word is a consonant (as opposed to a vowel) (Guy 1980; Bybee 2002; Raymond et al. 2006; Raymond et al. 2016).

Additionally, aspects of the greater discourse context can influence target word articulations in predictable ways. The variant forms of words and sounds are shaped in production by factors such as repeated mentions (Fowler and Housum 1987; Pisoni 1997; Baker and Bradlow 2009: 394), priming (Poplack 1980, 1984; Poplack and Tagliamonte 1989; Scherre and Naro 1991, 1992; Scherre 2001; Tamminga 2016), target speaking rate (Gahl 2008; Cohen Priva 2015; Lohmann 2018), and pragmatic function (Plug 2010). Stylistic variation and extralinguistic factors specific to the participants, such as gender, dialect, age, and social networks are also factors influencing the phonetic shape that words and sounds take (Labov 1972, 1994). The effects of factors such as these, and many more, are fairly well documented.

These numerous linguistic and extralinguistic factors of the phonetic, discourse and social context of speech operate simultaneously, in coordination and in opposition, to shape distinct realizations of words. However, importantly, this production variability, in turn, affects the lexical representations of word forms in memory (Bybee 1999; Hay 2018). As speakers experience the production and perception of words, episodic traces of such experiences are captured in memory (Pierrehumbert 2001, 2006). In this regard, words' proportions of use in conditioning contexts are crucial to understanding pronunciation variation (Bybee 2002). In usage-based accounts of language a consideration of words' context of use, or usage history, is crucial because the opportunity biases yield probabilistic phonetic outputs that accrue in memory. Contexts of use thus probabilistically determine the strongest and/or most numerous exemplars and, as such, give rise to variant forms stored in memory.

The phonetic shapes of words registered in the minds of speakers, which reflect the accumulation of contextual pressures, may then serve as targets for subsequent articulations. Given this type of feedback loop (Kemmer and Barlow 2000:ix) whereby experienced forms are stored and then later selected as targets for production, variation evident in targets may reflect factors of the immediate (online) context as well as the long-term histories of the words in these contexts. That is to say, in models of linguistic variation, we would predict evidence of both online contextual pressures, as well as cumulative effects of words' usage patterns.

Indeed, studies have demonstrated this to be the case for deletion and lenition of segments. For instance, in the case of word-final t/d deletion in English, Raymond et al. (2016) establish that, in addition to the most important conditioning factor for deletion being present in the target context (i.e., that the following word begins with a consonant), words' deletion rates reflect historical likelihood of use in this deleting context. Similarly, in a study of word-initial /f-/ variation in Spanish, Brown and Alba (2017) show that rates of fricative reduction reflect both the phonetic context of the moment of target realization and the target word's history of use in the phonetic context for reduction. Such findings, as predicted by Bybee (2002), reflect speakers' accumulated experience with words in specific phonetic environments. These studies and others like them (Brown 2004; Brown 2009; Brown and Raymond 2012; Raymond and Brown 2012; Brown 2015; Forrest 2017) measure the cumulative impact of extra-lexical phonetic context on reduction of sounds at word boundaries.

Other studies, nevertheless, measure the cumulative impact of conditioning factors on a different type of lenition: durational shortening of words or segments. For example, Gahl (2008) and Lohmann (2018) used a word's prior probability, or its lexical frequency, to predict durational shortening. Words of higher frequency are more likely to be pronounced with a more accelerated speech rate than words of lower frequency. This is likewise true of homophone pairs, so that the higher frequency member (e.g., *time*) is pronounced with a shorter duration than the lower frequency member (e.g., *thyme*). The cumulative effect of lexical frequency is evident while controlling for the online conditioning factors, such as speech rate of the target word context.

In addition to lexical frequency, informativity, or the average predictability of a word from its lexical context, also shapes word and segment duration. Seyfarth (2014) examines acoustic durations of content words in two corpora of spoken English to test whether speakers' prior experiences with words that vary in degree of informativity influence spoken word durations. Seyfarth (2014) finds that acoustic durations of words reflect the factors operating in the target context (online effects), such as local predictability. Words with low informativity are on average more predictable from lexical context and thus likely to be given a reduced articulation by speakers, regardless of their predictability in context (a lexical effect). Likewise, Cohen Priva (2015) finds that a segment's informativity, or its average local predictability from segmental contexts, predicts the segment's duration. Less informative segments have shorter durations in their local contexts, and low informativity segments have shorter durations even when word length, phonetic properties (such as speech rate), frequency, and local predictability are controlled. Thus, multiple studies modeling durational differences in production (Gahl 2008; Seyfarth 2014; Cohen Priva 2015; Lohmann 2018) report effects of words' prior distributions in critical contexts (implicating lexical representations) as well as expected effects of the online context.

Figuring prominently among the local predictors of target durations in reduction studies is the production rate of a target's speech context, either preceding or following the target. In fact, contextual speech rate is nearly universally considered in studies of phonological variation due in large part to the strength of its conditioning effect. With regard to durational shortening of target forms, speech rate is one of the strongest predictors for segments (Brown and Raymond 2014; Fosler-Lussier and Morgan 1999; Hualde and Prieto 2014; Nadeu 2014) and words (Arnon and Cohen Priva 2013). As the speech rate of the target context increases, target durations decrease.

Nevertheless, any long-term, cumulative effect of usage patterns in fast vs. slow speech has not been fully considered. Because faster speech contexts promote durational shortening of words (Gahl 2008, 2012; Seyfarth 2014; Lohmann 2018) and segments (File-Muriel and Brown 2011; Cohen Priva 2015), there is evidence that a word that occurs more often in fast speech may, on average, be more reduced than a word that usually occurs in slow speech, independent of the contextual rate in which that word is produced (Brown and Raymond 2014). This work is the first of its kind to test for an effect of evidence of words' cumulative likelihood of use in fast (vs. slow) speech on word and segment duration.

**2. Current Study**

To test for an effect of a word's usage history in the conditioning environment of fast speech, we chose a widely studied example of variation for which local, or online, contextual conditioning is well understood: /s/ variation in Spanish. We conduct acoustic analyses of segment and word durations in sociolinguistic interviews of the Spanish of Cali, Colombia (File-Muriel and Brown 2011). In this project we have two primary research questions. We examine what effect, if any, a word form's ratio of conditioning (FRC – Forms' Ratio of Conditioning) in fast speech contexts ($FRC_{RATE}$)–operationalized as the probability that a word occurs in fast speech–has on the speech rate at which a target word is produced. That is, does each word's cumulative exposure to fast speech predict online target rates? Additionally, this work explores whether this cumulative measure (FRC) will also predict the duration of a segment within a word. Results show that independent of multiple control factors known to constrain word and segment production rates, words' histories of conditioning in relatively fast/slow speaking rates predict target production durations. Words used proportionally more often in speaker-relative fast contexts are produced more quickly than those words used less often in accelerated speech rate contexts. We interpret these results as indicative of lexicalized effects of articulation rate in support of an exemplar model of lexical representation (Bybee 2001).

**3. Data**

To explore the possibility of words' ratio of occurrence in fast speech modulating target durations (an $FRC_{RATE}$ effect), this work examines variable productions of /s/ in Spanish (orthographic "c" before {"i", "e"}, "s", and "z") in addition to durations of the words containing the segment /s/.

Spanish /s/ variation has been extensively researched. Traditionally, the various realizations of /s/ have been classified auditorily and impressionistically into three discrete categories (Guitart 1976; Terrell 1977, 1979; Cedergren 1978; Lipski 1985). As exemplified in our data, these categories reflect maintained, sibilant articulations [s]/[z] (as in Example 1), aspirated articulations realized as a voiceless, glottal fricative [h] (Example 2), or deleted (Ø) segments (Example 3).

(1) *la univer*[s]*idad no* 'the university doesn't' [CAL-F-04]

(2) *las i*[h]*las del* 'the islands of' [CAL-F-09]

(3) *ahí tre*[Ø] *veces* 'there three times' [CAL-F-01]

More recent work has eschewed discrete categories of variants in favor of acoustic analyses reflecting gradience and a continuum of realizations (e.g. Erker 2010; File-Muriel and Brown 2011; Ryant and Liberman 2016). Variant acoustic realizations of /s/ in Spanish varieties include measures of duration of /s/, degree of /s/ intervocalic voicing (Torreira and Ernestus 2012; Hualde and Prieto 2014), forms and timing of post-aspiration of /s/ (Parrell 2012; Ruch 2013), and a continuum in strength of the spectral centroid or Center of Gravity of /s/ (File-Muriel 2012; Brown et al. 2014).

To investigate the effect of $FRC_{RATE}$ on the duration of /s/ and the duration of words containing /s/, speech was analyzed from eight native speakers of Spanish from and living in Cali, Colombia, collected in sociolinguistic interviews (File-Muriel and Brown 2011). The interviews were conducted by an in-group speaker of Caleño Spanish, a speaker who herself was born in Cali and raised in and around the city. Social variability was reduced by restricting our analysis to female college students in order to focus on the usage-based variables.

Studies of Spanish /s/ variation classically restrict focus to syllable-final position (Resnick 1975:13). Nevertheless, reduction is also found in syllable-initial position in multiple varieties of Spanish (Espinosa 1909; Flórez 1951; Sánchez 1982; López Scott 1983; Lipski 1984, 1986; Cotton and Sharp 2001; Brown and Torres Cacoullos 2003; Brown 2005, Brown 2009), including the variety of Spanish under analysis in this work (Brown and Brown 2012). Thus, from the transcripts of each recorded conversation, we targeted all fluently produced words spoken by our speakers that contained (at least one) /s/ in minimally five connected minutes of speech starting at minute 10:00 of each recording.

The target word types in the selected intervals included both content and function words found in these data of naturalistic, recorded conversations. The data extracted from the conversations represented 1,067 /s/-word types containing 4,652 tokens of /s/ in all syllable (onset, coda) and word (initial, medial, final) positions. Note that some /s/- word types contained more than a single /s/ segment (e.g., *islas* 'islands', *precisamente* 'precisely'; N = 228 types, 609 tokens). Additionally, some tokens of /s/ were morphemic (e.g. verbal marker 2ps, plural marker), but these account for a subset of tokens found solely in word-final position. Although Plag, Lohmann, Hedia and Zimmermann (2019) report durational differences of /s/ in English according to morphological category, many studies of word-final /s/ reduction in Spanish fail to find support for a role of morphemic status in variant realizations of /s/ (Poplack 1980; Hundley 1987; Ranson 1993; Labov 1994; Samper Padilla 2003; Ruiz Sanchez 2005; Schwegler and Kempff 2007).

Each token of /s/ was manually delimited with time-stamped interval boundaries in TextGrids using the acoustic software Praat (Boersma and Weenink 2019). Additionally, the word in which /s/ occurred was delimited, as was the pause-bounded Intonation Unit (IU) in which each target word was produced. After completing the manual delimitation of /s/, words with /s/, and intonational unit boundaries, a Praat script automated the extraction of the durations in milliseconds of /s/ targets, /s/-words, and the IUs used in this study. This procedure enabled a calculation of the local speech context rate following each target and each speaker's average global speaking rates (calculated across a speaker's IUs, excluding the target word), as well as /s/ durations and target word production rates. All speech rates were calculated in phones per second, as described below.

Excluded from the data were two /s/-word tokens of the English word 'jet ski' and all tokens of *entonces* 'so, therefore, thus' (N = 113). The latter presented extreme, perhaps lexicalized, variation as a discourse marker (see, for example, Travis, 2005), including, for example, four instances of *entonces* realized as [ton]. Also excluded were the /s/-words containing /s/ tokens (N = 2) with production rates more than two standard deviations above the speakers' mean speaking rate (using zero-centered and scaled z-score rate values). Two tokens of /s/ words produced with outlying rates greater than 20 phones per second (two tokens of *es*) were excluded. Finally, because we are interested in the effects of following speech context on production, we excluded all tokens of /s/ words followed by a pause (i.e., IU final).

## 4. Methods

### *4.1. Data partitioning and data subsets.*

Prior to modeling word and /s/ rates, the data were partitioned into two subsets. The goal of data partitioning was to enable independence of the predictors (post-target context rates and words' likelihood of use in speaker-relative fast speech) and the dependent variables regarding word rates and /s/ durations. One subset (referred to as the calculation set) was used to create control factors entered in the models (such as a speaker's average speaking rate), as well as our variable of interest, $FRC_{RATE}$. The other subset (the test set) was used to analyze the effects of the independent variables on the dependent variables (/s/ duration and /s/-word production rate). The steps used to create the two sets are as follows.

An approximation of the likelihood that a word occurs in fast speech was estimated as the number of times a word occurs in fast speech out of the total number of times it occurs in a corpus. This estimate obviously becomes more reliable as the number of word tokens from which it is calculated increases, but requiring a larger number of tokens also decreases the number of data items. We chose four occurrences of a word in our overall dataset as the cutoff, allowing two occurrences with which to calculate likelihood of occurrence in fast speech and two tokens for inclusion in the model data. Thus, the /s/-tokens in all words for which we had more than four word tokens in the data (with at least two tokens to estimate $FRC_{RATE}$) participated in data partitioning. The /s/-tokens in all word types for which we had fewer than four tokens in the overall dataset (N = 848) were assigned to the calculation set prior to partitioning. The /s/-tokens in words with four or more tokens in the data were sorted by speaker, word type, and /s/-timestamp. The /s/ tokens were then assigned by parity to the test set and to the calculation set, with the odd numbered tokens assigned to the calculation set and the even numbered tokens assigned to the test set. This procedure produced an approximately even split of /s/ tokens and types per speaker into both sets. These steps yielded 381word types comprising 2,428 /s/ tokens in the calculation set and 166 word types comprising 1,375 /s/ tokens in the test set. Note that the exact number of tokens in both models was further restricted, as detailed below.

Average speaking rates were then calculated for each speaker using the calculation set data. A speaker's mean speaking rate was estimated by averaging post-word context rates in the calculation set, measured in phones per second from the right side of token to the end of the intonational unit (cf. Gahl 2008).[i] Speaker average post-target speaking rates varied, as can be seen in Table 1.

Table 1: Speaker Average Post-context Rates (Phones per Second)

| **Speaker** | **Mean** | **Standard deviation** | **Tokens** |
|---|---|---|---|
| 1 | 10.6 | 5.4 | 145 |
| 2 | 14.5 | 7.1 | 172 |
| 3 | 10.2 | 5.8 | 862 |
| 4 | 9.6 | 6.9 | 181 |
| 5 | 11.6 | 7.1 | 183 |
| 6 | 12.3 | 6.2 | 132 |
| 7 | 11.4 | 6.8 | 363 |
| 8 | 13.2 | 7.2 | 389 |

The speaker average post-target speaking rates were used to calculate each word type's likelihood of occurring in fast post-context speech relative to its speaker's mean post-context rate ($FRC_{RATE}$). For each /s/-word token it was determined whether the token's post-context rate was faster or slower than the average post-context rate for that speaker (from the calculation subset). Next, using each token's classification as being used in a relatively fast or slow target context by speaker, a ratio was calculated for the word type. The number of times each word type occurred in speech that was faster than the average for the speaker who spoke it out of all the instances of that word type was calculated as a ratio between 0 and 1. Thus, a ratio of 0 signifies that the word type never occurred in speech that was faster than the average for the speaker who spoke it, and a ratio of 1 indicates that all the tokens of the word type were spoken in speech that was faster than the average for speakers. For example, the word type *centro* 'center' was spoken a total of 5 times by 3 different speakers. In one instance, *centro* was produced in a context that was faster than average for its speaker. In the remaining four instances, *centro* was produced in a context that was slower than average for its speaker. Thus, the word type *centro* has a likelihood of occurrence in fast speech that can be estimated as 1/5 = 0.2. Across all word types the mean likelihood of occurrence in fast speech is 0.403 (s.d. = 0.254; median = 0.444). Our $FRC_{RATE}$ variable, thus, is a predictability measure, estimating how likely fast speech is for a given word. Note, however, that the variable included in our models, as well as other probability measures, was modified as described in §6.1.

The likelihood of occurrence in fast speech values estimated from the calculation subset for each word type were used to code $FRC_{RATE}$ for the equivalent word tokens in the test subset. Note that through this procedure the likelihood of occurrence of each word in fast speech, a continuous measure used as a predictor in the test subset, was independent of the /s/-word post-context rates, /s/-token production durations, and /s/-word production rates of the test tokens.

**5. Dependent factors**

To study the rate at which /s/-words were produced, all /s/-word production rates in the test set were normalized to be used as a dependent variable in our word rate and /s/ duration models. Centered and scaled z-scores were created per speaker by subtracting from each /s/-word token's rate in phones per second the mean word rate (across all words) of its speaker in phones per second and then dividing those values by the standard deviation of the speaker's word rates. The resulting normalized word rate values were centered around zero (0) and reflect speakers' degree of speech rate deviation from mean. Values above 0 represent word rates higher than the mean word rate for that speaker (fast target words) and values below 0 represent word rates lower than the mean word rate for the speaker (slow target word).

In modeling /s/-word rate, word tokens containing multiple target tokens of /s/ (for example, *después* 'after'), were filtered such that only one occurrence of each target word was analyzed in the word rate model. However, both tokens of /s/ in such a word were analyzed in the subsequent /s/ duration model. Other exclusions in the word rate and /s/ duration models include target words followed by pauses and targets preceded or followed by word recyclings and word cutoffs. These methods produced 1,221 /s/-word tokens (from 159 types) for analysis of word duration.

To study the production durations of the /s/ segment, the durations of non-deleted /s/-tokens were log-transformed and then normalized by subtracting from each log /s/ duration its speaker's average log /s/ duration (excluding deleted [s] phones) and dividing by the speaker's standard deviation of the non-deleted log /s/ durations. The resulting values are centered around zero (0), with z-scores above 0 representing tokens of /s/ with a longer duration than the average for the speaker who spoke it. These methods gave us 1,024 tokens of /s/ (152 word types) for analysis of /s/ duration.

Table 2: Summary of dependent variables

| **Dependent variable** | **Continuous Measures (zero centered scaled z-scores)** |
|---|---|
| Word rate: | Difference from speaker mean speaking rate (phones per second) / standard deviation of speakers' speaking rate |
| /s/ duration: | Difference of log /s/ duration from speaker mean log /s/ duration / standard deviation of speakers' log of /s/ durations |

## 6. Independent factors

### *6.1. Word-rate model.*

The rate at which words are articulated is a function of several predictor variables, both linguistic and social. In addition to our variable of interest, a word's likelihood of occurrence in fast post-context speech ($FRC_{RATE}$), we control for multiple conditioning factors of the target context. For each target word we coded the speech rate of the post /s/-word context, the length of the /s/-word in phones (as pronounced by the speaker), the /s/-word's grammatical class (noun, verb, modifier, function word), the lexical frequency of the /s/-word, the predictability of the target word from the word following the target (Pr(w|w+1)), and the predictability of the target word from the word preceding the target (Pr(w|w-1)), as described below.

Contextual speech rates fluctuate and constrain duration and speech rate of target forms (Gahl 2008; Bell et al. 2009). Average post-context rate per speaker was calculated using the post-contexts of target words in the calculation subset. The local, contextual speech rate for each target word in the test set was normalized as a difference from the speaker's mean speaking rate. Normalized speaking rates of the context following target words with a positive value were spoken faster than average (in phones per second) for that speaker (fast contexts); negative values represent post-context rates slower than average for that speaker (slow contexts).

Word rates vary as a factor of the number of sounds in the word (Terrell 1977, 1979; File-Muriel 2010). A positive correlation exists between the time it takes to articulate a word and the length of that word in phones. We coded for the number of canonical phones in target words. Instances in which a word evidenced deletion of /s/ were coded for the number of sounds pronounced. For example, if *tus* 'your' (pos. adj.) was pronounced without the coda /s/, the word token was coded as having two phones [tu], as opposed to three.

Additionally, it has been argued that the duration of words reflects the status of the word as either being a function word or content word (Jurafsky et al. 2001; Bell et al. 2002). Owing to processing differences or mechanisms in speech production models, function words are uttered more quickly than content words, as they tend to carry less meaning, are less informative, and are often more predictable from the surrounding context. Recent work (Lohmann and Conwell 2020) attributes acoustic differences in duration of words to grammatical classes beyond the binary categorization of function vs. content word. As such, all token /s/ words were coded into four grammatical classes: (1) function words (such as *los* 'masculine singular definite article 'the'', *nos* 'first person plural object 'us, we, ourselves', *se* '3rd singular pronoun'); (2) nouns and pronouns (e.g., *autobús* 'bus', *ellos* 'they - 3 p. p.'); verb forms (e.g., *dice* '3rd singular says'); and modifiers (adjectives and adverbs, such as *social* 'social', *antes* 'before').

Lexical frequency is also a known correlate of speaking rates, such that higher frequency words correlate with faster speaking rates and are durationally shorter (Jurafsky et al. 2001; Bell et al. 2002; Aylett and Turk 2004). This effect is thought to reflect processing differences–a result of ease of lexical retrieval given the relatively higher activation of frequent words and/or articulatory reduction indicative of an automation of articulatory gestures through practice (Bybee 1999). In addition to production ease, high frequency words are thought to be less informative or more predictable (Jaeger and Buz 2017), and accordingly are given reduced articulations. Word frequency was calculated for word lexemes as opposed to word lemmas. That is, paradigmatically related forms (*estás, estamos, estuviera,* etc. of *estar* 'to be') are not conflated, but rather are each accorded unique frequency values. We chose to calculate frequencies for word types (and not lemmas) because previous work on /s/ variation (Brown et al. 2014) finds that type frequency better predicts variation compared to lemma frequency.

Word and segment rates

To produce the maximum likelihood estimates (MLEs) for our word probability variables (word frequency and word predictabilities), we used the OpenSubtitles (Lison and Tiedemann 2016) Spanish corpus (OPUS) of films (1,169,065,308 words). Corpus counts were then adjusted using Laplace smoothing (Brysbaert and Diependaele 2012; Cohen Priva and Jaeger 2018), because some target words in our dataset were not in the OPUS corpus. The OPUS corpus was also used to calculate the conditional probability of the target word given the immediately following word (Pr(w|w+1)) and preceding word (Pr(w|w-1)). Probabilistic reduction is widely reported, with increased predictabilities of words and segments correlating with increased phonetic reduction (Jurafsky et al 2001; Raymond et al. 2006). We again used Laplace smoothing, which allowed predictability estimates of word bigrams in our overall dataset that were not found in the OPUS corpus.

For our variable of interest–the predictability of fast following speech contexts given a /s/ word ($FRC_{RATE}$)–we used counts of target words' occurrences in fast and slow speech in the calculation dataset, as described in §4.1. There were 29 that never occurred in fast speech; once again Laplace smoothing was employed to produce MLEs for these types. However, the estimates made using counts in the calculation dataset are poor for low-frequency words (i.e., words for which we only have 2 tokens in the training data with which to estimate $FRC_{RATE}$). These values are mostly 0 (or sometimes 1 = or 0.5). To better estimate the distribution of these low frequency words, we used the method employed in Baayen and Sproat (1996) for estimating the distribution of a word form in alternating contexts. Following this procedure, we calculated the rate of occurrence in fast speech of all hapax forms in the calculation set (i.e., all /s/ words in the calculation data that only occurred once, e.g., *bastante* 'enough, a lot', which occurs only once, in slow speech; and *cierta* 'sort of, kind of', which occurs only once, in fast speech). This average is .38. We use this value for the MLE of all words in the test data whose $FRC_{RATE}$ score was based on only two tokens in the calculation set.

Finally, all probability variables (word frequency, Pr(w|w+1), Pr(w|w-1), and $FRC_{RATE}$) were converted into bits of information by a negative $\log_2$ transformation.

*6.2. /s/ duration model*

The second research question we explore in this work is whether the cumulative rate of occurrence in a conditioning context (in this case fast speech context) impacts the production of segments within words. In modeling /s/-durations within the words, we include our variable of interest ($FRC_{RATE}$), as well as other factors known to constrain target phone durations and Spanish /s/ articulations in particular. These control variables included word length as pronounced (that is, the number of segments realized), normalized speech rate of the context following the target /s/ word, lexical frequency, conditional probability of the following word, and grammatical category of the /s/ word, (as described for word durations in §6.1). Additionally, /s/ tokens were coded for preceding and following phonetic context, described in the following.

Previous research has shown that adjacent phones significantly constrain the duration of /s/ in Spanish (File-Muriel and Brown 2011; Brown et al. 2014; Hualde and Prieto 2014; Henriksen and Harper 2016, Strycharczuk and Kohlberger 2016). We were interested in testing whether one cumulative measure ($FRC_{RATE}$) could significantly predict duration across the entirety of the word and thus include all word and syllable positions of /s/ [word-initial (N = 204); word-medial, syllable-initial (N = 351); word-medial, syllable-final (N = 105), word-final (N = 373)]. Target tokens of /s/ in these Spanish words, thus, are adjacent to multiple phonetic contexts. Preceding phonological context was coded as pause (IU initial) or other (any consonant or vowel segment). Preceding pauses have been shown to condition longer duration of sounds including /s/ (File-Muriel and Brown 2011) compared to other segmental contexts. The following phonetic contexts were coded as following /t/, other segment (consonant or vowel; recall that no target words in the data were followed by a pause). Previous work (Méndez Dosuna 1985; Brown 2009) finds the duration of the fricative /s/ is significantly longer before /t/.

Table 3: Summary of control factors

| Control factor | Word rate model | /s/ duration model |
|---|---|---|
| Log likelihood of use in fast post-context speech ($FRC_{RATE}$) | ✔ | ✔ |
| Normalized speech rate of the post /s/-word context | ✔ | ✔ |
| Length of the /s/-word in phones (as pronounced) | ✔ | ✔ |
| Grammatical class of the /s/-word (noun, verb, modifier, function word) | ✔ | ✔ |
| Log lexical frequency of the /s/-word | ✓ | ✔ |
| Log of target word predictability from following word (*w* \| *w*+1) | ✔ | ✔ |
| Log of target word predictability from preceding word (*w* \| *w*-1) | ✔ | ✔ |
| Phonetic context preceding /s/ (pause, no-pause) | | ✔ |
| Phonetic context following /s/ (/t/, no-/t/) | | ✔ |

Each dependent variable (i.e., word rate, /s/ duration) was examined in a mixed effect statistical model using lmer function with the lme4 R package (R Development Core Team 2011). Additionally, in both mixed effect statistical models, random intercepts were entered for speaker and word, and random slopes were entered for speaker; however, the word rate model failed to converge with random slopes included, so these were remove and the model was run without these in the final model. Following Gries 2013 (259-261), we perform backward variable selection of factors starting with a model including all main effects and their pairwise interactions. Factors and interactions contributing least significantly to the model at each step are iteratively removed, resulting in the final models we report.

## 7. Results

### *7.1. Word-rate model*

We submitted the 1,221 tokens of /s/ words (159 types) from our data set to a linear mixed effect model predicting normalized word rates (faster or slower than speaker average) with speaker and word type as random effects. The p values are determined via the Satterthwaite's degrees of freedom method. We opt for a speaker normalized value over an absolute duration value. Cohen Priva and Gleason (2018: 1512) note that although speakers differ in their average speaking rates, speech rates of individuals persist rather consistently across production contexts. An accelerated speaking rate for one individual might correspond to a reduced rate for another. In perceiving speech as either fast or slow, therefore, listeners must normalize by speaker in forming these judgments. The results of model (1) are summarized in Table 4.

(1) [Speaker normalized word rate ~ (1 | speaker) + (1 | token) + log(Pr(w|w+1)) + logPr(*w* | *w* - 1)) + following context speaking rate + word length in phones as

pronounced + log(word frequency) + grammatical class + log($FRC_{RATE}$) + context speaking rate:grammatical class].

Table 4: Mixed Effects Linear Regression Predicting Speech Rate of Spanish Word with /s/

| **Random effects** | **Variance** | **Std. Dev.** | | | | |
|---|---|---|---|---|---|---|
| Word (intercept) | 0.361044 | 0.60087 | | | | |
| Speaker (intercept) | 0.003494 | 0.05911 | | | | |
| Residual | 4.103708 | 2.02576 | | | | |
| **Fixed effects**[ii] | **Estimate coef.** | **Std. error** | **t-statistic** | **df** | **p-value** | **Sig.** |
| (Intercept) | 1.72087 | 0.41531 | 4.144 | 129.06142 | 6.14e-05 | *** |
| -$\text{Log}_2$ ($FRC_{RATE}$) | -0.52396 | 0.19907 | -2.632 | 211.80515 | 0.009113 | ** |
| -$\text{Log}_2$ Pr(*w*\|*w* +1) | -0.06210 | 0.02063 | -3.010 | 983.41153 | 0.002682 | ** |
| -$\text{Log}_2$ Pr(*w*\|*w* -1) | -0.06362 | 0.02229 | -2.854 | 1131.62498 | 0.004393 | ** |
| -$\text{Log}_2$(word frequency) | 0.0.01728 | 0.03884 | 0.445 | 299.84390 | 0.656629 | ns |
| Following speech context rate (normalized) | 0.59429 | 0.12675 | 4.689 | 1199.13774 | 3.06e-06 | *** |
| Word length as pronounced | 0.31656 | 0.05342 | 5.926 | 189.65186 | 1.44e-08 | *** |
| Lex. class = function | | | | | | |
| Adverbials | -1.88389 | 0.31623 | -2.795 | 106.12689 | 0.006159 | ** |
| Nouns / pronouns | -0.77627 | 0.29837 | -2.602 | 108.80821 | 0.010566 | * |
| Verbs | -0.57652 | 0.30048 | -1.919 | 113.50369 | 0.057541 | . |
| Speech rate x Adverbial | -0.62860 | 0.177758 | -3.540 | 1138.97868 | 0.000417 | *** |
| Speech rate x Nouns / pronouns | -0.19469 | 0.17031 | -1.143 | 1204.60275 | 0.253204 | ns |
| Speech rate x Verbs | -0.22872 | 0.17324 | -1.320 | 12000.55521 | 0.187004 | ns |

N = 1221 Random effects N: Word = 159, Speaker = 8
AIC = 5390.016
Significance codes: p-value = 0 *** 0.001 ** 0.01 * 0.05 . 0.1 ns 1

Word and segment rates

The effect of the local speech rate was, as expected, significant, with relatively accelerated post-context rates predicting increased target word speech rates, as has been reported in previous studies. Moreover, the rate at which a word is spoken was significantly predicted by its likelihood of occurrence in fast versus slow speech (our cumulative measure, $FRC_{RATE}$). Words with a greater proportion of use in speaker-normalized relatively fast speech had increased rates of production over those used less often in fast speech. This cumulative effect was independent of the speech rate of the target context at production and supports previous findings that words' history of use in relatively fast speech accrues in memory in a lexically specific way (Brown and Raymond 2014).

The conditional probability of the target given the following word and the preceding word also significantly predict target word duration in these data. The /s/ words that are more predictable from the previous and following words were spoken more quickly than less predictable target words, possibly reflecting in part production planning effects (Tanner, Sonderegger and Wagner 2017). This result aligns with studies reporting increased reduction and speech rates in words that are more predictable (Jurafsky et al 2001; Raymond et al. 2006).

Target words with a greater number of phones were also spoken at a faster rate than words with fewer phones. This result is consistent with Spanish /s/-reduction studies that report greater reduction of /s/ in longer words (Terrell 1977, 1979; File-Muriel 2010). The grammatical class of the target word also significantly predicts word rate in these data. Adverbials and nouns (and verbs marginally so) are spoken at a significantly lower rate (i.e., slower) than function words. Grammatical class also has a significant interaction with speech rate. As context speaking rate increases, we find quicker word rates in each grammatical class with the exception of modifiers.

Unlike previous research on /s/ reduction (File-Muriel 2009; File-Muriel and Brown 2011; Brown et al. 2014), we do not find a significant effect of word frequency on the rate at which words are spoken in this study. We suspect that some potential effect of word frequency may be captured through correlations with other predictors in the model such as grammatical class (function words, for example, generally being more frequent) or word length (shorter words typically being more frequent).

In summary, aspects of the word, including length in phones, grammatical class and cumulative measures such as $FRC_{RATE}$ as well as aspects of the target (online) context such as post-context speaking rate and word predictabilities together predict the rate at which the target words are produced. Importantly, the rate at which the content and function words containing /s/ were spoken in the corpus of conversational Colombian Spanish responds to the words' distribution in relatively fast speech. Word that occurred proportionally more often in fast speech (high $FRC_{RATE}$) were produced more quickly than words that occurred proportionally less often in this conditioning environment. The correlation between our variable of interest ($FRC_{RATE}$) expressed as bits of information and word duration is seen in Figure 1.

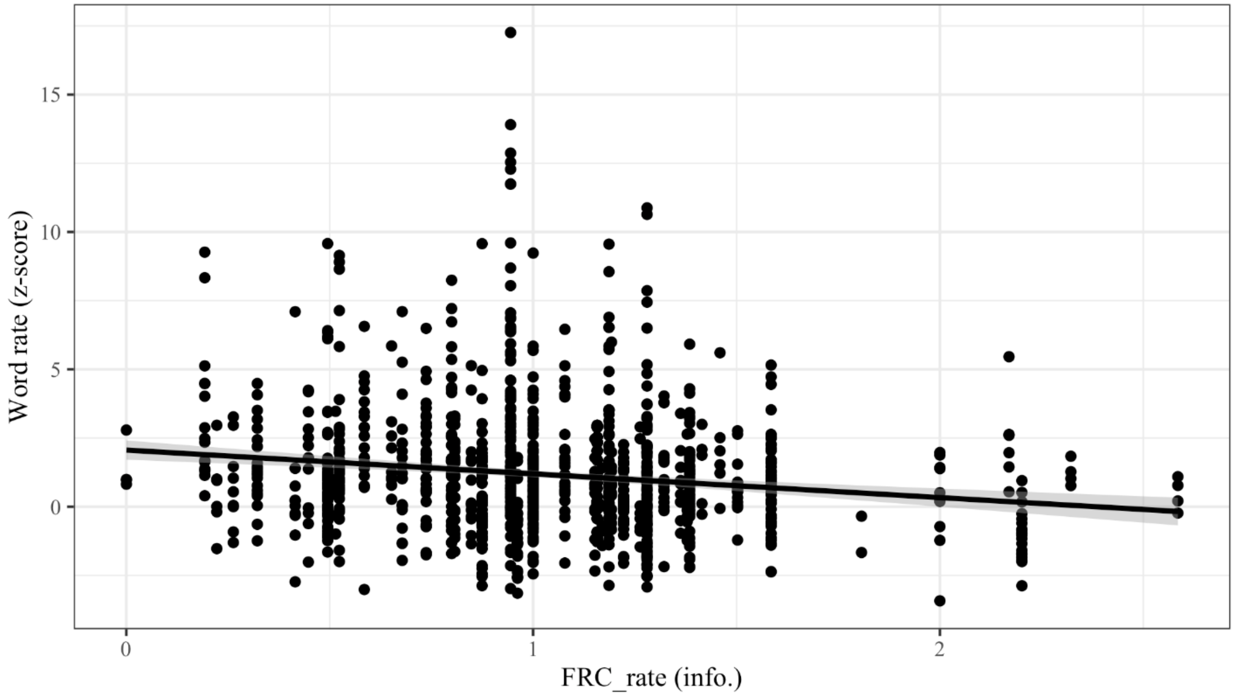


Figure 1: Speaker-normalized speech rate of Spanish /s/ words by word Forms' Ratio of Conditioning in speaker-relative fast speech (information bits)

This newly identified $FRC_{RATE}$ effect is independent of factors known to correlate with speech rate, such as target context rate, word frequency, word predictabilities and word length.

*7.2. /s/ duration model*

For our model predicting /s/ duration, we removed from the test set the instances of /s/ that were deleted in production. We submitted the remaining 1024 non-deleted tokens of /s/ (152 types) from our data set to a linear mixed effect model predicting normalized durations of the /s/ segments (/s/ durations shorter or longer than average per speaker) with word type entered as random intercept and random intercepts and random slopes for speaker with our variable of interest, $FRC_{RATE}$. The p-values are determined via the Satterthwaite's degrees of freedom method. The results of the model in (2) are summarized in Table 5.

Word and segment rates

(2) [Log of speaker normalized /s/ duration ~ (1 + $FRC_{RATE}$ | speaker) + (1 | token) + $FRC_{RATE}$ + log(*w* | *w* + 1) + log(*w* | *w* - 1) + context speaking rate + log(word frequency) + grammatical class + previous phone + following phone + word length as pronounced + log(w|*w* – 1):word length as pronounced + context speaking rate:previous phone + context speaking rate:following phone + log(word frequency):grammatical class + log(word frequency):word length as pronounced].

Table 5: Mixed Effects Linear Regression Predicting the Duration of Spanish /s/

| **Random effects** | **Variance** | **Std. Dev.** | | | | |
|---|---|---|---|---|---|---|
| Word (intercept) | 0.0000000 | 0.00000 | | | | |
| Speaker (intercept) | 0.0239490 | 0.15475 | | | | |
| Speaker (slope $FRC_{RATE}$) | 0.0006643 | 0.02577 | | | | |
| Residual | 0.5282840 | 0.72683 | | | | |
| **Fixed effects**[iii] | **Estimate coef.** | **Std. error** | **t-statistic** | **df** | **p-value** | **Sig.** |
| (Intercept) | 3.624e-01 | 3.393e-01 | 1.068 | 7.115e+02 | 0.28585 | |
| $-Log_2$ ($FRC_{RATE}$) | 1.838e-01 | 6.114e-02 | 3.006 | 9.782e_01 | 0.00336 | ** |
| $-Log_2$ Pr(*w*\|*w* +1) | 2.110e-02 | 7.458e-03 | 2.829 | 1.000e_03 | 0.00476 | ** |
| $-Log_2$ Pr(*w*\|*w* -1) | 3.737e-02 | 1.804e-02 | 2.072 | 9.991e+02 | 0.03855 | * |
| $-Log_2$(word frequency) | -1.433e-01 | 3.651e-02 | -3.925 | 1.002e+3 | 9.26e-05 | *** |
| Following speech context rate (normalized) | -1.933e-01 | 2.578e-02 | -7.499 | 1.000e+3 | 1.41e-13 | *** |
| Post /t/ = yes | 5.020e-01 | 7.824e-02 | 6.416 | 9.988e+02 | 2.16e-10 | *** |
| Pre. phon. context = pause | 3.770e-01 | 1.485e-01 | 2.538 | 9.990e+02 | 0.01129 | * |
| Lex class = function | | | | | | |
| Adverbial | 6.249e-01 | 3.424e-01 | 1.825 | 1.002e+03 | 0.06829 | . |
| Noun / pronouns | -1.168e-01 | 3.381e-01 | -0.346 | 1.003e+03 | 0.72979 | n.s. |

| | | | | | | |
|---|---|---|---|---|---|---|
| Verb | -7.166e-01 | 3.382e-01 | -2.119 | 1.000e+03 | 0.03437 | * |
| Verb | -7.166e-01 | 3.382e-01 | -2.119 | 1.000e+03 | 0.03437 | * |
| Word length as pronounced | -7.894e-02 | 4.713e-02 | -1.675 | 1.000e+03 | 0.09428 | . |
| -Log$_2$ Pr($w$\|$w$ +1) x Word length | -7.245e-03 | 3.135e-03 | -2.311 | 9.994e+02 | 0.02104 | * |
| Speech rate x Prec. Phonology | 3.243e-01 | 1.262e-01 | 2.570 | 9.977e+02 | 0.01031 | * |
| Speech rate x Post /t/ | 1.952e-01 | 6.453e-02 | 3.025 | 1.001e+03 | 0.00255 | ** |
| -Log$_2$(word frequency) x modifier | -1.781e-02 | 3.140e-02 | -0.569 | 1.002e+03 | 0.56929 | n.s. |
| -Log$_2$(word frequency) x noun | 3.327e-02 | 3.269e-02 | 1.082 | 1.003e+03 | 0.27948 | n.s. |
| -Log$_2$(word frequency) x verb | 8.450e-02 | 3.279e-02 | 2.577 | 1.000e_03 | 0.01010 | * |
| -Log$_2$(word frequency) x Word length | 1.189e-02 | 4.014e-03 | 2.963 | 1.001e+03 | 0.00312 | ** |

N = 1024; Random effects N: Word = 152, Speaker = 8
AIC = 2401.695
Significance codes: p-value = 0 *** 0.001 ** 0.01 * 0.05 . 0.1 ns 1

The results of a mixed effects linear regression show that our cumulative measure ($FRC_{RATE}$) has an effect on the normalized duration of /s/, such that as $FRC_{RATE}$ increases, duration of /s/ decreases. That is, tokens of /s/ are durationally shorter in words with greater likelihood of occurrence followed by accelerated speech (i.e. speech that is quicker than the average for each speaker in this study). Put another way, the cumulative exposure to fast speech yields shorter /s/ productions, even when local speaking rate is statistically controlled.

Although we do not include deleted tokens of /s/ in this model, the average $FRC_{RATE}$ for deleted /s/ tokens is significantly higher (.57) than the average $FRC_{RATE}$ for non-deleted /s/ segments (.50) [t = 6.042, df = 356.8, p-value = 3.829e-09]. The higher ratio of occurrence in relatively fast speech for the deleted segments compared to maintained /s/ segments is in the direction we would expect given that fast speech conditions many types of reduction, including deletions. This cumulative effect reflects in our data the fact that words with deleted /s/ segments occur in faster speech compared to non-deleted /s/ segments (avg. normalized del .097 and non-del .065).

As expected, local speaking rate has a significant effect on the normalized duration of /s/: tokens of /s/ are shorter when followed by speech that is quicker than the average for each speaker, similar to what is reported in File-Muriel and Brown (2011). Other predictor variables also significantly condition the duration of /s/. The following phonological context influences its duration, such that /s/ is shorter when followed by a sound other than /t/, that is, a following /t/ increases the duration of /s/. This phonetic conditioning is also reported in previous studies (Mendez Dosuna 1985). The preceding phonological context also conditions /s/ duration: /s/ is longer when preceded by a pause. Tokens of /s/ contained within function words (*se* '3rd p.s. impersonal/reflexive pronoun', *nos* '1st p.p. object pronoun', *además* 'also', etc.) are significantly shorter than /s/ tokens within adverbials and nouns, but not significantly shorter than /s/ in verbs (*sabe* '3rd p.s. knows', *dice* '3rd p.s. says', *hablamos* '1st p.p. speak/talk'). The predictabilities of the target (given the following word and given the preceding word) had a significant effect on predicting duration of tokens of /s/ within the target. As predictability of the /s/ word increases, duration of the /s/ in the target decreases.

The lexical frequency of the word containing /s/ also significantly predict /s/ duration within these words, although not in a direction we might have predicted. As information content goes up (and frequencies go down) we would expect segment durations to go up. The effect of lexical frequency on /s/ in Spanish, nevertheless, has not been consistently reported in previous research (see, for example, the discussion in Brown et al. 2014).

There are two significant interactions in the data between the target contextual speaking rate and the phonetic contexts of the /s/. The durational shortening effect of a faster post-target context speaking rate on tokens of /s/ is nullified when the preceding phonetic context is a pause, and when the following phonetic context is a /t/. In both of these phonetic contexts which disfavor durational shortening of /s/, the durational shortening effect of the target context is neutralized.

Lexical frequency interacts with grammatical class and word length. We find that frequency exerts an influence in the /s/ tokens found in verbs, but not in the other grammatical categories (function words, nouns, modifiers). For length, we find that for shorter words, /s/ tokens in high frequency words have a longer duration than /s/ in low frequency words. In longer words, there is no difference in duration predicted by the frequency. Lastly, the interaction with word length and word predictability reflects a greater effect of predictability for shorter words compared to longer words. The effect on /s/ duration from the predictability of the target given the preceding word decreases as words get longer.

In sum, the duration of /s/ segments in the Spanish content and function words reflects each word's history of occurrence in relatively fast speech. The /s/ segments in word forms with a higher ratio of conditioning by fast speech (high $FRC_{RATE}$) are durationally shorter than /s/ segments in words with a lower ratio of occurrence in the conditioning environment. This effect on /s/ duration of our variable of interest expressed on an information scale is illustrated in Figure 2.

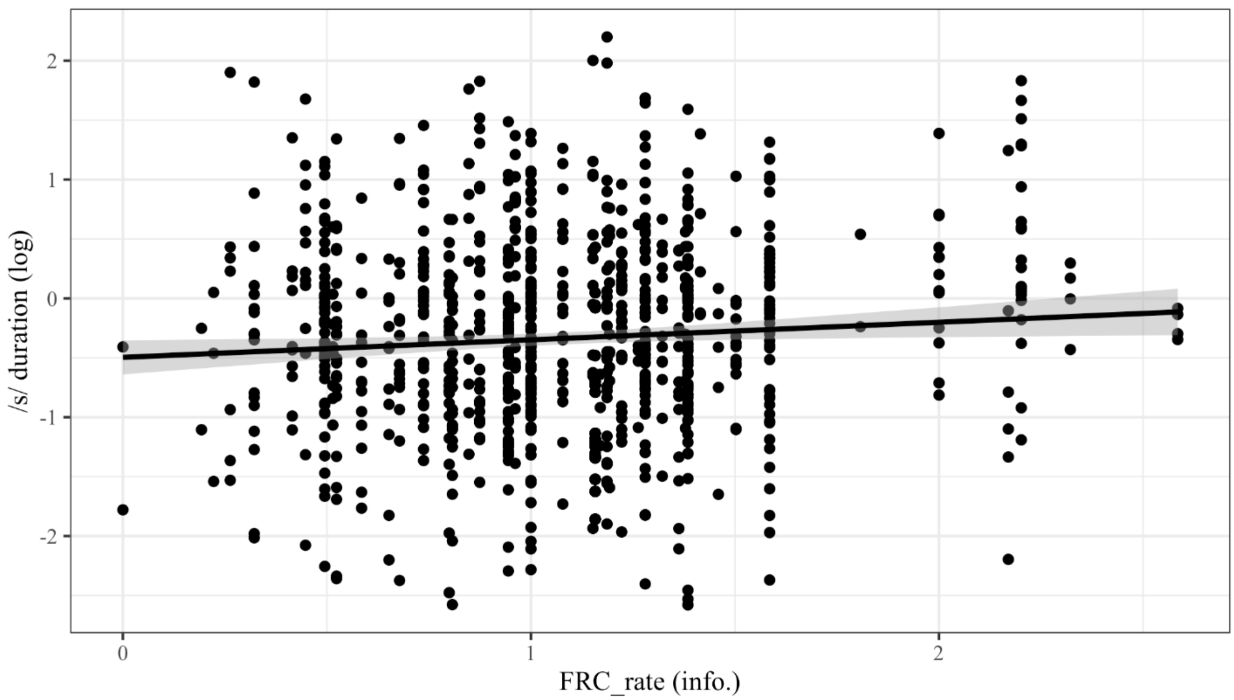


Figure 2: Speaker-normalized duration of Spanish /s/ by word Forms' Ratio of Conditioning in speaker-relative fast speech (information scale)

This $FRC_{RATE}$ effect is independent of factors known to predict /s/ duration, such as target speaking rate, phonetic context and position in the word.

**8. Discussion and conclusions**

This work is the first of its kind to explore impacts on words and segments of patterns of use in relatively fast speech contexts. In a corpus of naturalistic spoken Spanish, we estimated a word's proportion of use in contexts conditioned by relatively fast speech (which is equivalent to the predictability of fast speech given that word) and applied this measure ($FRC_{RATE}$) to a widely studied case of variation (variable durations of /s/ and speaking rates of words containing /s/ in Spanish). Results show that the cumulative measure significantly predicts both the rate at which the word is spoken (Table 4) as well as the duration of the /s/ segments in words (Table 5). As the target forms' ratio of use conditioned by fast speech increases, the rate of production increases for the word and segment, independent of the factors present in the targets' context. That is, higher $FRC_{RATE}$ predicts durationally shorter words and /s/ segments (Figures 1 and 2). This result is consistent with studies reporting other cumulative effects (e.g., Seyfarth 2014; Foulkes and Hay 2015; Soskuthy and Hay 2017), in that we report a lexically specific effect that reflects outcomes of the cumulative conditioning from words' usage patterns.

We interpret this result of our work as evidence in support of an Exemplar Model of lexical representation (Bybee 2001). As words appear in speaker-relative fast speech contexts, their likelihood of speeded articulation increases. Rapid realizations yield shorter durations across the word and across individual segments within the word. The phonetic outputs of these experiences in production and perception form part of the episodic traces of these words in memory. Aspects of the phonetic form registered in memory, thus, are not just limited to deletions, lenitions, or the like. Duration (or speaking rate) of a word is reflected in its lexical representation and is not the sole purview of the enactment of an articulation plan. The accumulated effects of the conditioning environment (relatively fast context rates) is evident in our results in conjunction with the online effect of the targets' conditioning environment. Thus, this work adds following speech context rate to the list of contextual conditioning factors that leave an imprint on the lexicon through a form's ratio of conditioning (FRC).

The measure we employ in this work ($FRC_{RATE}$) is different from previous cumulative measures employed in studies of Spanish /s/. Previous research of /s/ variation has found a significant FRC effect (Brown 2004; Brown 2009; Raymond and Brown 2012) that measured the likelihood of a word occurring in phonetic reducing environments in adjacent words on realizations of /s/ at word boundaries. These works estimate each word's ratio of occurrence in an extra-lexical context that phonetically conditions lenition of /s/. For word-initial /s/, for example (Brown 2004; Raymond and Brown 2012), each words' likelihood of occurrence in a phonetic environment conditioning lenition (a preceding non-high vowel) predicts reduction. For a word-final /s/, each word form's ratio of conditioning by a following consonant-initial word (Brown 2009) predicts reduction. In each case, those works report an effect of /s/ segments at word boundaries that persists independently of the phonetic context in which the target word is measured. The current work employs a measure predicting segment realizations (duration) of targets not limited to those found at word boundaries.

Effects of the post-context conditioning are evident for all tokens of /s/ in the words in our data, and our cumulative measure is thus able to capture contextual effects on segments not at word boundaries. That is, the measure we present can reach inside and across words to predict an aspect of segment realization not just for segments at word boundaries, but also for segments in word-medial positions.

This result perhaps suggests avenues of inquiry for other word-medial processes. For instance, any medial variants conditioned by speech rate may well reflect cumulative echoes of word forms' ratio of conditioning in fast speech (as the /s/ segments in this work do). Lexical accumulation of speech rate effects might be evident for word-medial /t,d/ deletion in English (e.g.; Raymond et al. 2006), especially given that context rate itself has an effect on deletion. Similarly, other processes, such as voiced approximant lenition in Spanish (e.g.; Vasilescu, Hernandez, Vieru and Lamel 2018), or vocalic examples such as medial schwa deletion (eg., *memory* vs. *mammary*) in English (Bybee 2001) might also be predicted by such a measure.

Previous work (e.g. Baese-Berk, Dilley, Vinke and Banzina 2019) has demonstrated that (contextual and distal) speech rate plays an important role in speech perception. As Baese-Berk et. al (2019) note, in addition to employing semantic and syntactic likelihoods in spoken word and segment recognition, perhaps "listeners also utilize speech rate to make predictions" (589). This study suggests that in addition to perceptual effects, speech rate plays an important role (online, cumulatively) in the variant forms of words that are produced. As such, the probabilistic reduction correlated with factors of word predictability given semantic and syntactic context (Jaeger and Buz 2017: 48-49), may also include speech rate likelihood.

It is noteworthy that in this study *word frequency* did not significantly predict production rates of Spanish /s/ words, despite the importance ascribed to this linguistic predictor in adjudicating between models of lexical representation or in ranking contributions of (extra)linguistic factors to variation and change. As Bayley, Greer and Holland (2017: 434) note, frequency can be like a "Cheshire cat, appearing fully in some studies, faintly in others, and not at all in still others". We do not take a lack of a significant frequency effect (such as the one in our word rate model Table 4) as evidence counter to exemplar models (Bybee 2001, 2010). Instead, we suggest that if different, more multi-dimensional types of frequency are foregrounded (Raymond and Brown 2012; Gries 2012:5; Divjak 2018:19) we can obtain a better picture of the variation. To improve models of linguistic variation and change, future projects should incorporate measures that are more appropriate for capturing the complexities of experiences with words, including contextualized probability distributions.

**Endnotes**:

[i] Note that for this and all speech rate calculations diphthongs were treated as 1.5 phones, rather than 2 phones, given their intermediary duration between singly-occurring vowels on the one hand and vowels in hiatus on the other hand (cf. Aguilar 1999, Aguilar & Machuca 1995). Thus, the phrase *la ciudad* was counted as having 7.5 phones, and not 8 phones.
[ii] Note that the probability measures, the first four independent factors listed in the table, were entered as bits of information. Negative correlations thus indicate decreasing word rates correlated with decreasing probabilities.
[iii] The positive correlation equates to durational shortening for the four predictability measures expressed as 'bits of information' ($FRC_{RATE}$, word predictabilities, word frequency).

**References**:

Aguilar, Lourdes. 1999. Hiatus and diphthong: Acoustic cues and speech situation differences. *Speech Communication 28*(1). 57-74.

Aguilar, Lourdes & María Machuca. 1995. Pragmatic factors affecting the phonetic properties of diphthongs. *Fourth European Conference on Speech Communication and Technology*, EUROSPEECH 1995, Madrid, Spain, September 18-21.

Arnon, Inbal & Uriel Cohen Priva. 2013. More than words: The effect of multi-word frequency and constituency on phonetic duration. *Language and Speech 56*(3). 349-371.

Aylett, Matthew & Alice Turk. 2004. The smooth signal redundancy hypothesis: A functional explanation for relationships between redundancy, prosodic prominence, and duration in spontaneous speech. *Language and Speech 47*(1). 31-56.

Baayen, Harald & Richard Sproat. 1996. Estimating lexical priors for low-frequency morphologically ambiguous forms. *Computational Linguistics 22*(2). 155-166.

Baese-Berk, Melissa M., Laura C. Dilley, Molly J. Henry, Louis Vinke & Elina Banzina. 2019. Not just a function of function words: Distal speech rate influences perception of prosodically weak syllables. *Attention, Perception, & Psychophysics 81*(2). 571-589.

Baker, Rachel & Ann Bradlow. 2009. Variability in word duration as a function of probability, speech style, and prosody. *Language and Speech 52*(4). 391-413.

Bayley, Robert, Kristen Greer & Cory Holland. 2017. Lexical frequency and morphosyntactic variation. *Spanish in Context 14*(3). 413-439.

Bell, Alan, J.M. Brenier, Michelle Gregory, Cynthia Girand & Daniel Jurafsky. 2009. Predictability effects on durations of content and function words in conversational English. *Journal of Memory and Language 60*(1). 92-111.

Bell, Alan, Michelle Gregory, J.M. Brenier, Daniel Jurafsky, Ayako Ikeno & Cynthia Girand. 2002. Which predictability measures affect content word duration. In *PMLA. Proc. ISCA Workshop on Pronunciation Modeling and Lexicon Adaptation for Spoken Language, Estes Park, CO* (pp. 1-5).

Boersma, Paul & David Weenink. 2019. Praat: doing phonetics by computer (Version 6.0. 52) [Windows].

Brown, Earl K. 2009. *A Usage-Based account of syllable-and word-final /s/ reduction in four dialects of Spanish.* Munich: Lincom Europa.

Brown, Earl K. & Matthew Alba. 2017. The role of contextual frequency in the articulation of initial /f/ in Modern Spanish: The same effect as in the reduction of Latin /f/?. *Language Variation and Change 29*(1). 57.

Brown, Earl K. & Esther L. Brown. 2012. "Syllable-final and syllable-initial /s/ reduction in Cali, Colombia: One variable or two?" In Richard File-Muriel & Rafael Orozco (eds.), *Linguistic Studies in Colombian Varieties of Spanish*, 89-106, Madrid: Iberoamericana.

Brown, Earl K., Michael Gradoville & Richard File-Muriel. 2014. The variable effect of form and lemma frequencies on phonetic variation: Evidence from/s/realization in two varieties of Colombian Spanish. *Corpus Linguistics and Linguistic Theory 10*(2). 213-241.

Brown, Esther L. 2004. *The reduction of syllable-initial/s/in the Spanish of New Mexico and southern Colorado: A usage-based approach*. The University of New Mexico dissertation.

Brown, Esther L. 2005. New Mexican Spanish: Insight into the variable reduction of" la ehe inihial"(/s-/). *Hispania*. 813-824.

Brown, Esther L. 2015. The role of discourse context frequency in phonological variation: A usage-based approach to bilingual speech production. *International Journal of Bilingualism 19*(4). 387-406.

Brown, Esther L. & Rena Torres Cacoullos. 2003. Spanish [s]: A different story from beginning (initial) to end (final). In Rafael Navarro, Luis López, & Richard Cameron (eds.), *A Romance perspective on language knowledge and use: Selected papers from the 31st Linguistic Symposium on Romance Languages,* 21-38. Chicago: John Benjamins.

Brown, Esther L. & William D. Raymond. 2012. How discourse context shapes the lexicon: Explaining the distribution of Spanish f-/h-words. *Diachronica 29*(2). 139-161.

Brown, Esther & William D. Raymond. 2014. Contextual frequency effects in Spanish phonology. Paper presented at the Georgetown University Round Table. Georgetown University, Washington, DC. March 14.

Brysbaert, Marc & Kevin Diependaele. 2013. Dealing with zero word frequencies: A review of the existing rules of thumb and a suggestion for an evidence-based choice. *Behavior Research Methods 45*(2). 422–430. https://doi.org/10.3758/s13428-012-0270-5.

Bybee, Joan. 1999. Usage-based phonology. In Kathleen Wheatley, Michael Darnell, Michael Noonan, Frederick J. Newmeyer and Edith A. Moravcsik (eds.), *Functionalism and formalism in linguistics: Volume I: General papers*, 211-242. Philadelphia: John Benjamins Publishing Company.

Bybee, Joan. 2001. *Phonology and language use*. Cambridge, MA: Cambridge University Press.

Bybee, Joan. 2002. Word frequency and context of use in the lexical diffusion of phonetically conditioned sound change. *Language Variation and Change 14*(3). 261-290.

Bybee, Joan. 2010. *Language, usage, and cognition*. Cambridge, MA: Cambridge University Press.

Cedergren, Henrietta J. 1978. En torno a la variación de la s final de sílaba en Panamá: Análisis cuantitativo. In Humbero Lopez Morales (ed.), *Corrientes actuales en la dialectología del Caribe hispánico*, 37–49. Puerto Rico: Editorial Universitaria.

Cohen Priva, Uriel. 2015. Informativity affects consonant duration and deletion rates. *Laboratory Phonology 6*(2). 243-278.

Cohen Priva, Uriel & Emily Gleason. 2018. The role of fast speech in sound change. Paper presented at the Annual Conference of the Cognitive Science Society. Madison, WI. July, 2018.

Cohen Priva, Uriel & T. Florian Jaeger. 2018. The interdependence of frequency, predictability, and informativity in the segmental domain. *Linguistics Vanguard 4*(S2).

Cotton, Eleanor & John Sharp. 2001. *Spanish in the Americas*. Washington, D.C.: Georgetown University Press.

Divjak, Dagmar. 2017). The role of lexical frequency in the acceptability of syntactic variants: Evidence from *that*-clauses in Polish. *Cognitive Science 41*(2). 354-382.

Erker, Daniel. 2010. A subsegmental approach to coda /s/ weakening in Dominican Spanish. *International Journal of the Sociology of Language.* (203). 9-26.

Espinosa, Aurelio. 1909. *Studies in New Mexican Spanish: Part I: Phonology* (No. 53). Albuquerque, N.M.: University of New Mexico Press.

File-Muriel, Richard. 2009. The role of lexical frequency in the weakening of syllable-final lexical /s/ in the Spanish of Barranquilla, Colombia. *Hispania*. 348-360.

File-Muriel, Richard. 2010. Lexical frequency as a scalar variable in explaining variation. *The Canadian Journal of Linguistics 55*. 1-25.

File-Muriel, Richard. 2012. A laboratory approach to s-lenition in the Spanish of Barranquilla, Colombia. In Richard File-Muriel & Rafael Orozco (eds.), 127-140. *Linguistic studies in Colombian varieties of Spanish*. Madrid: Iberoamericana.

File-Muriel, Richard & Earl K. Brown. 2011. The gradient nature of s-lenition in Caleño Spanish. *Language Variation and Change 23*(2). 223-243.

Flórez, Luis. 1951. *La pronunciación del español en Bogotá* (Vol. 8). Bogotá, Colombia: Instituto Caro y Cuervo.

Forrest, Jon. 2017. The dynamic interaction between lexical and contextual frequency: A case study of (ING). *Language Variation and Change 29*(2). 129-156.

Fosler-Lussier, Eric & Nelson Morgan. 1999. Effects of speaking rate and word frequency on pronunciations in conversational speech. *Speech Communications 29*. 137-158.

Foulkes, Paul & Jennifer Hay. 2015. The emergence of sociophonetic structure. In Brian MacWhinney and William O'Grady (eds), 292. *The handbook of language emergence*. Hoboken: Wiley-Blackwell.

Fowler, Carol & Jonathan Housum. 1987. Talkers' signaling of “new” and “old” words in speech and listeners’ perception and use of the distinction. *Journal of Memory and Language 26*(5). 489-504.

Gahl, Susanne. 2008. Time and thyme are not homophones: The effect of lemma frequency on word durations in spontaneous speech. *Language 84*(3). 474-496.

Gahl, Susanne. 2012. Why so short? Competing explanations for variation. In *Proceedings of the 29th West Coast Conference on Formal Linguistics* (Vol. 29, pp. 1-10). Cascadilla Proceedings Project, Somerville, MA.

Gries, Stefan Th. 2012. Introduction. In Stefan Th. Gries & Dagmar Divjak (eds.), *Frequency effects in language learning and processing*, 1-6. The Hague: Mouton de Gruyter.

Gries, Stefan Th. 2013. *Statistics for linguistics with R: A practical introduction*. Berlin: Walter de Gruyter.

Guitart, Jorge M. 1976. *Markedness and a Cuban dialect of Spanish*. Washington, DC: Georgetown University Press.

Guy, Gregory. 1980. Variation in the group and the individual: The case of final stop deletion. In W. Labov (ed.), *Locating language in time and space*, 1-36. New York: Academic Press.

Hay, Jennifer. 2018. Sociophonetics: The role of words, the role of context, and the role of words in context. *Topics in Cognitive Science 10*(4). 1-11.

Henriksen, Nicholas & Sarah Harper. 2016. Investigating lenition patterns in south-central Peninsular Spanish /sp st sk/ clusters. *Journal of the International Phonetic Association.* 1-24.

Hualde, José Ignacio & Pilar Prieto. 2014. Lenition of intervocalic alveolar fricatives in Catalan and Spanish. *Phonetica 71*(2). 109-127.

Hundley, James E. 1987. Functional constraints on plural marker deletion in Peruvian Spanish. *Hispania 70*(4). 891-894.

Jaeger, T. Florian & Esteban Buz. 2018. Signal reduction and linguistic encoding. In Eva Fernández & Helen Smith Cairns (eds.), *Blackwell handbooks in linguistics. The handbook of psycholinguistics*, 38–81. Hoboken, NJ: Jon Wiley & Sons.

Jurafsky, Daniel, Alan Bell, Michelle Gregory & William D. Raymond. 2001. The effect of language model probability on pronunciation reduction. Paper presented at the International Conference on Acoustics, Speech, and Signal Processing (ICASSP) 2001, Salt Lake City, UT.

Kemmer, Suzanne & Michael Barlow. 2000. Introduction: A usage-based conception of language. In Michael Barlow & Suzanne Kemmer (eds.), *Usage-based models of language,* vii-1. Stanford: CSLI.

Labov, William. 1972. *Sociolinguistic patterns*. Philadelphia, PA: University of Pennsylvania Press.

Labov, William. 1994. *Principles of linguistic change, Volume 1: Internal factors.* Oxford: Blackwell.

Lipski, John. 1984. On the weakening of /s/ in Latin American Spanish. *Zeitschrift für Dialektologie und Linguistik*. 31-43.

Lipski, John. 1985. /s/ in Central American Spanish. *Hispania 68*. 143–149.

Lipski, John. 1986. Instability and reduction of /s/ in the Spanish of Honduras. *Revista canadiense de estudios hispánicos*. 27-47.

Lison, Pierre & Jörg Tiedemann. 2016. *OpenSubtitles2016: Extracting Large Parallel Corpora from Movie and TV Subtitles.* In Proceedings of the 10th International Conference on Language Resources and Evaluation (LREC 2016). Retrieved from: http://opus.nlpl.eu/OpenSubtitles-v2018.php.

Lohmann, Arne. (2018). Time and thyme are not homophones: A closer look at Gahl's work on the lemma-frequency effect, including a reanalysis. *Language 94*(2). e180-e190.

Lohmann, Arne & Erin Conwell. 2020. Phonetic effects of grammatical category: How category-specific prosodic phrasing and lexical frequency impact the duration of nouns and verbs. *Journal of Phonetics 78*. 100939.

Lopez Scott, Alma Leticia. 1983. *A sociolinguistic analysis of/s/ variation in Honduran Spanish* University of Minnesota dissertation.

Méndez Dosuna, Julian. 1985. La duración de S en los grupos SP, ST, SK: A propósito del orden regular de difusión en algunos cambios fonéticos. In José L. Melena (ed.), *Symbolae Ludovico Mitxelena Septuagenario Oblatae, 647-655.* Vitoria: Universidad del Pais Vasco.

Nadeu, M. (2014). Stress-and speech rate-induced vowel quality variation in Catalan and Spanish. *Journal of Phonetics 46*, 1-22.

Parrell, Benjamin. 2012. The role of gestural phasing in Western Andalusian Spanish aspiration. *Journal of Phonetics 40*(1). 37-45.

Pierrehumbert, Janet. 2001. Exemplar dynamics: Word frequency, lenition, and contrast. *Typological Studies in Language 45*. 137-158.

Pierrehumbert, Janet. 2006. The statistical basis of an unnatural alternation. *Laboratory Phonology 8*. 81-107.

Pisoni, David. 1997. Some thoughts on 'normalization' in speech perception. In Keith Johnson & J. W. Mullenix (eds.), *Talker variability in speech processing*, 9-32. San Diego, CA: Academic Press.

Plag, Ingo, Sonia Ben Hedia, Arne Lohmann & Julia Zimmermann. 2019. In Livia Körtvélyessy & Pavel Stekauer (eds.), *Complex Words,* 260-292. Cambridge: Cambridge University Press.

Plug, Leendert. 2010. Pragmatic constraints in Usage-Based Phonology, with reference to some Dutch phrases. *Journal of Pragmatics* (42). 2014-35.

Poplack, Shana. 1980. The notion of the plural in Puerto Rican Spanish: Competing constraints on (s) deletion. *Locating language in time and space. 1*. 55-67.

Poplack, Shana. 1984. Variable concord and sentential plural marking in Puerto Rican Spanish. *Hispanic Review 52*(2). 205-222.

Poplack, Shana & Sali Tagliamonte. 1989. There's no tense like the present: Verbal-s inflection in early Black English. *Language Variation and Change 1*(1). 47-84.

R Development Core Team (2011). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL http://www.R-project.org/.

Ranson, Diana L. 1993. The interaction of linguistic and contextual number markers in Andalusian Spanish. *Hispania 76*. 919-930.

Raymond, William D., Robin Dautricourt & Elizabeth Hume. 2006. Word-internal /t, d/ deletion in spontaneous speech: Modeling the effects of extra-linguistic, lexical, and phonological factors. *Language Variation and Change 18*(1). 55.

Raymond, William D. & Esther L. Brown. 2012. Are effects of word frequency effects of context of use? An analysis of initial fricative reduction in Spanish. In Stefan Th. Gries & Dagmar Divjak (eds.), *Frequency effects in language learning and processing, 35-52.* The Hague: Mouton de Gruyter.

Resnick, Melvin. 1975. *Phonological variants and dialect identification in Latin American Spanish.* The Hague: Mouton.

Ruch, Hanna. 2013. Investigating a gradual metathesis: Phonetic and lexical factors on/s/-aspiration in Andalusian Spanish. *University of Pennsylvania Working Papers in Linguistics 19*(2). 19.

Ruíz Sánchez, Carmen. 2005. Restricciones funcionales en la elisión del marcador de plural en el español de Caracas. *Boletín de Lingüística 17*. 89-105.

Ryant, Neville & Mark Liberman. 2016. Large-scale analysis of Spanish/s/-lenition using audiobooks. In *Proceedings of Meetings on Acoustics 22ICA* (Vol. 28, No. 1, p. 060005). Acoustical Society of America.

Samper Padilla, José Antonio. 2003. La variación fonológica: los estudios hispánicos sobre -/s/ implosiva. Accessed at http://cvc.cervantes.es/obref/congresos/valladolid/ ponencias/unidad_diversidad_d el_espanol/1_la_norma_hispanica/samper_j.htm.

Sánchez, Rosaura. 1982. Our linguistic and social context. In Jon Amastae & Elías Lucía Olivares (eds.), *Spanish in the United States: Sociolinguistic aspects*, 9-46. Cambridge: Cambridge University Press.

Scherre, Maria Marta Pereira. 2001. Phrase level parallelism effect on Noun Phrase number agreement. *Language Variation and Change 13*. 91–107.

Scherre, María Marta Pereira & Anthony Naro. 1991. Marking in discourse: “Birds of a feather”. *Language Variation and Change 3*(1). 23-32.

Scherre, María Marta Pereira & Anthony Naro. 1992. The serial effect on internal and external variables. *Language Variation and Change 4*(1). 1-13.

Schwegler, Armin and Juergen Kempff. 2007. *Fonética y fonología españolas*, 3rd edn. New York: John Wiley & Sons, Inc.

Seyfarth, Scott. 2014. Word informativity influences acoustic duration: Effects of contextual predictability on lexical representation. *Cognition 133*(1). 140-155.

Sóskuthy, Martón & Jennifer Hay. 2017. Changing word usage predicts changing word durations in New Zealand English. *Cognition 166*. 298-313.

Strycharczuk, Patrycja & Martin Kohlberger. 2016. Resyllabification reconsidered: On the durational properties of word-final/s/in Spanish. *Laboratory Phonology 7*. 24.

Tagliamonte, Sali A. 2012. *Variationist sociolinguistics. Change, observation, interpretation*. Malden, MA / Oxford: Wiley-Blackwell.

Tamminga, Meredith. 2016. Persistence in phonological and morphological variation. *Language Variation and Change 28*(3). 335-356.

Tanner, James, Morgan Sonderegger & Michael Wagner. 2017. Production planning and coronal stop deletion in spontaneous speech. *Laboratory Phonology: Journal of the Association for Laboratory Phonology 8*(1). 15.

Terrell, Tracy. 1977. Constraints on the aspiration and deletion of final/s/in Cuban and Puerto Rican Spanish. *Bilingual Review/La Revista Bilingüe 4*(1/2). 35-51.

Terrell, Tracy. 1979. Final /s/ in Cuban Spanish. *Hispania 62*(4). 599-612.

Torreira, Francisco & Mirjam Ernestus. 2012. Weakening of intervocalic/s/in the Nijmegen Corpus of Casual Spanish. *Phonetica 69*(3). 124-148.

Travis, Catherine. 2005. *Discourse Markers in Colombian Spanish: A Study in Polysemy*. Berlin/New York: Mouton de Gruyter.

Vasilescu, Ioana, Nidia Hernandez, Bianca Vieru & Lori Lamel. 2018. Exploring temporal reduction in dialectal Spanish: A large-scale study of lenition of voiced stops and coda-s. *Interspeech*. 2728-2732.

Word and segment rates