UNIVERSITY OF MURCIA

International Journal of English Studies

IJES

www.um.es/ijes

Measures of speech rhythm and the role of corpus-based word frequency: a multifactorial comparison of Spanish(-English) speakers

MICHAEL J. HARRIS
STEFAN TH. GRIES*
University of California, Santa Barbara

Received: 9 October 2011 / Accepted: 3 December 2011

ABSTRACT
In this study, we address various measures that have been employed to distinguish between syllable and stress-timed languages. This study differs from all previous ones by (i) exploring and comparing multiple metrics within a quantitative *and* multifactorial perspective and by (ii) also documenting the impact of corpus-based word frequency. We begin with the basic distinctions of speech rhythms, dealing with the differences between syllable-timed languages and stress-timed languages and several methods that have been used to attempt to distinguish between the two. We then describe how these metrics were used in the current study comparing the speech rhythms of Mexican Spanish speakers and bilingual English/Spanish speakers (speakers born to Mexican parents in California). More specifically, we evaluate how well various metrics of vowel duration variability as well as the so far understudied factor of corpus-based frequency allow to classify speakers as monolingual or bilingual. A binary logistic regression identifies several main effects and interactions. Most importantly, our results call the utility of a particular rhythm metric, the PVI, into question and indicate that corpus data in the form of lemma frequencies interact with two metrics of durational variability, suggesting that durational variability metrics should ideally be studied in conjunction with corpus-based frequency data.

KEYWORDS:
Speech rhythm, syllable-timed vs. stress-timed, PVI, Spanish, English, monolingual vs. bilingual, corpus frequencies

RESUMEN
In this study, we address various measures that have been employed to distinguish between syllable and stress-timed languages. This study differs from all previous ones by (i) exploring and comparing multiple metrics within a quantitative *and* multifactorial perspective and by (ii) also documenting the impact of corpus-based word frequency. We begin with the basic distinctions of speech rhythms, dealing with the differences between syllable-timed languages and stress-timed languages and several methods that have been used to attempt to distinguish between the two. We then describe how these metrics were used in the current study comparing the speech rhythms of Mexican Spanish speakers and bilingual English/Spanish speakers (speakers born to Mexican

* ***Address for correspondence***: Michael J. Harris and Stefan Th. Gries, University of California, Santa Barbara, CA. USA. E-mail: <michaelharris@umail.ucsb.edu> and <stgries@linguistics.ucsb.edu>.
The order of authors is arbitrary.

© Servicio de Publicaciones. Universidad de Murcia. All rights reserved.

ISSN: 1578-7044

parents in California). More specifically, we evaluate how well various metrics of vowel duration variability as well as the so far understudied factor of corpus-based frequency allow to classify speakers as monolingual or bilingual. A binary logistic regression identifies several main effects and interactions. Most importantly, our results call the utility of a particular rhythm metric, the PVI, into question and indicate that corpus data in the form of lemma frequencies interact with two metrics of durational variability, suggesting that durational variability metrics should ideally be studied in conjunction with corpus-based frequency data.

PALABRAS CLAVE:
Ritmo del habla, acompasamiento silábico vs. acompasamiento acentual, PVI, Español, Inglés, monolingüe y bilingüe, frecuencias de corpus.

1. INTRODUCTION

1.1. General introduction

Pike (1945) described the simple rhythm units of languages: stress-timed and syllable-timed. The former means that the durations of syllables vary according to the placement of stress and seem to be less uniform throughout the entire phrase; languages that sound more stress-timed include, for example, Dutch and English. The latter means that syllable duration is relatively uniform within a phrase, as in, say, French and Spanish.

For quite some time, this perceived difference in speech timing was regarded as a dichotomy: "As far as is known, every language in the world is spoken with one kind of rhythm or with the other" (Abercrombie 1967:97). Dasher & Bolinger (1982) suggested that stress-timing and syllable-timing can be correlated to specific phonological phenomena in a given language. Specifically, they suggested that (i) stress-timed languages present a greater variety of syllable types, and (ii) stress-timed languages have greater vowel reduction between stress and unstressed vowels. Dauer (1987) agreed that speech rhythms are a product of their phonological properties, but added that languages should not be thought of as either stress-timed or syllable-timed, but instead should be conceived on a continuum, with the most stress-timed languages at one extreme of the continuum and the most syllable-timed languages at the other extreme. While some languages – e.g., Spanish and English – exist near the opposite ends of this continuum, other languages exist somewhere in the middle, exhibiting some syllable-timed characteristics and some stress-timed characteristics. Catalan, for instance, has simple syllable types but also exhibits vowel reduction, causing it to fall somewhere between stress-timed and syllable-timed on the continuum. Dauer (1987) also concluded that simple measures of interstress intervals or syllable durations were not effective in assigning rhythm class, demonstrating the necessity of a measure of speech rhythms with more discriminatory power.

As a result of a more continuous view of speech rhythms, linguist sought to classify language rhythms relatively. That is, rather than classify a language as either syllable-timed or stress-timed, languages or varieties/dialects of languages were compared to other languages or other varieties/dialects of the same language. Low & Grabe (1995) made significant strides in the study of speech rhythms in their study of prosodic patterns in Singapore English compared to British English. This study is significant in focus and methodology. Firstly, it examined

© Servicio de Publicaciones. Universidad de Murcia. All rights reserved.

two varieties of the same language, comparing native (L1) and second language (L2) English speakers. Secondly, it introduced the Pairwise Variability Index (PVI, also sometimes called *r*PVI (*r* for *raw*) in later studies) as a new durational measure that is intended to relatively classify speech rhythms by measuring the variation between successive sets of vowels as opposed to the more limiting measure of overall variation within a phrase (as cited in Carter, 2007:5). This study served to provide a framework for further speech rhythm studies and studies of cross-varietal speech rhythms of a single language in particular.

The PVI is the measure of the absolute difference of the vowel duration of two adjacent syllables divided by the average vowel duration of the same two adjacent syllables (cf. equation below); thus, a sentence with *n* syllables yields *n*-1 PVIs. These PVIs represent the variability of vowel duration, with lower PVIs representing a more syllable-timed language and higher PVIs representing a more stress-timed language. Two things are important to note. First, PVIs do not represent an index of 'the absolute rhythm of a language' – instead, they allow the comparison of the rhythms of two or more languages or varieties. Second, Lows & Grabe's PVI is reported as a mean for each utterance (or even a speaker; cf. below), rather than a series of individual PVI scores.

Although the PVI was adopted by many linguists as an apparently transparent and accurate method of comparatively classification of speech rhythms, other metrics were also proposed under the heading of *interval measures* (or IMs, to use White & Mattys's term). Specifically, Ramus, Nespor, & Mehler (1999) defended the use of three variables in order to determine the rhythm of a language:

- the percentage of a sentence taken up by vowel duration (%V);
- the standard deviation of vowel duration (ΔV); and
- the standard deviation of consonant duration (ΔC).

Additionally, Ramus, Nespor, & Mehler suggested that ΔV and ΔC could be normalized for speech rate by dividing the metrics by the mean of the interval durations, which in effect yields variation coefficients.

Deterding (2001) proposed measuring the duration of the entire syllable rather than only the vowel duration, arguing that some syllables may be longer than others regardless of the presence of vowels. The author used a normalizing equation for syllable length similar to the PVI called the Variability Index (VI). However, this method was criticized by Thomas & Carter (2006:339) as too complicated and ambiguous in the practical application with the use of spectrograms, as syllable divisions prove difficult to define.

Low, Grabe, & Nolan (2000) revisited speech rhythms of Singapore and British English and discussed the application of speech-rate normalization to the PVI, the *n*PVI (normalized PVI) as opposed to the *r*PVI. This normalization, calculated by dividing the *r*PVI by the mean of durational differences is used by some linguists, while others, such as Carter (2005, 2007) and Thomas & Carter (2006) use the *r*PVI. While Low, Grabe, and Nolan (2000) compared first language (L1) speakers and second language (L2) speakers, the more recent studies of Carter (2005, 2007) as well as Fought (2003) investigated bilingual speakers

© Servicio de Publicaciones. Universidad de Murcia. All rights reserved.

of Spanish and English; more specifically, the participants in these latter studies were Chicanos who grew up speaking both Spanish and English. Also, in addition to using the *r*PVI, Thomas and Carter also reported a single mean PVI for each speaker, as opposed to Low and Grabe's computation of a mean for each utterance. Bunta and Ingram (2007) also examined the acquisition of Spanish-English bilingual speech rhythms in children using vocalic and intervocalic *n*PVI values, and concluded that bilingual and monolingual speakers differ in speech rhythms, particularly in the youngest participants, age approximately 4 years, although these differences diminished for older speakers.

White & Mattys (2007), undertook a more exhaustive investigation of the relative merits of the PVI and other IMs by applying both normalized and raw pairwise variability indices as well as the IMs suggested by Ramus, Nespor, & Mehler (1999) to both cross-varietal and cross-language utterances. Additionally, they tested the correlation of these metrics with speech-rate, exploring the effectiveness and/or necessity of rate-normalization measures. Although previous studies had compared competing metrics of durational variability (cf. Low, Grabe, & Nolan 2000), White & Mattys' (2007) study was more comprehensive: their goal was to compare the metrics used to classify durational variability, as opposed to the task of classifying the rhythms of specific languages or varieties. The study concluded that the rate-normalized *n*PVI, the rate-normalized standard deviation of vowel duration VarcoV (VARCOEFF in this study), as well as the percentage of an intonational unit comprised by vowels best distinguished between syllable and stress-timed languages as well as between L1 and L2 speakers.

Even though White & Mattys (2007) constitutes important progress, it still leaves room for improvement in several areas. First, the statistical assessment of the performance of the metrics is incomplete; they report whether or not certain effects are significant, but fail to report the relative strength of each effect. Thus, one cannot evaluate the relative statistical performance of each metric due to their failure to report R^2-values, beta coefficients, etc. Furthermore, their exploration does not include a truly multifactorial analysis, which could potentially reveal a far more complex picture of the behavior of the metrics. Secondly, a specific weakness of the PVI, is mentioned by the authors: "PVI scores derived alternating patterns and monotonic geometric series may be the same, so that, for example PVI(2, 4, 2, 4, 2, 4) and PVI(2, 4, 8, 16, 32, 64) are equal" (White & Mattys, 2007:519). The practice of reporting a mean PVI value for each utterance, speaker, or speaker type is statistically problematic; not only does the use of the means 'wash away' much information from the data as mentioned, but PVIs may not be normally-distributed (as they are in our data), which makes the mean a problematic measure of central tendency (see Section 4.6 below).

1.2. The present paper

The purpose of the present paper is twofold. First, we investigate the effect of bilingualism upon two groups of speakers: monolingual speakers of Mexican Spanish vs. Chicano bilingual speakers of American English and Mexican Spanish. These two groups of speakers make for an interesting test case given the opposing classifications of English and Spanish.

© Servicio de Publicaciones. Universidad de Murcia. All rights reserved.

The monolingual speakers 'only' speak what is traditionally regarded as a syllable-timed language whereas the bilingual speakers also speak what is traditionally regarded as a stress-timed language. It is not unreasonable to expect that this contrast will be reflected in the vowel duration variability of the speakers. It is this statistical effect that we target in this study, especially given the results of Carter (2005, 2007) and Fought (2003) who found that Chicano English tends to be more syllable timed (more similar to Spanish) than that of American English, and, in the case of Carter, African American English. Thus, it was expected that Chicano Spanish would be more stress-timed than that of their monolingual counterparts (cf. MacLeod & Stoel-Gammon 2005 for similar findings for voice-onset times in Canadian English and French). Table 1 visualizes the language-rhythm continuum as well as, in the last row, our overall assumption that the bilingual speakers will exhibit more of a stress-timed behavior; note that we only expect the bilingual speakers to exhibit more variable durations – we are not committing to a specific point on the continuum.

Poles of continuum	stress-timed ⟷			syllable-timed
Typical durations	variable durations			uniform durations
Language examples	Dutch, English			French, Spanish
Expected effect		bilingual speakers	>	monolingual speakers

Table 1: Observed and hypothesized relations between language and speaker types

Second, in comparing the mono- and bilingual speakers, we also compare the classificatory power of several language-rhythm measures. More specifically, we explore how much PVI and other durational measures such as standard deviations or variation coefficients can help distinguish between monolingual and bilingual speakers. By testing two varieties of the same language, we provide a rather simple case upon which to test the PVI as well as competing IMs. Some studies comparing metrics have included many more languages or varieties of languages, making interpretation of the results more convoluted and, potentially, ambiguous. Meanwhile, the vast majority of studies of durational variability as it relates to language-rhythm include far fewer speakers. White & Mattys (2007), for instance, only included one speaker for each language and/or variety. Although the current study's ten speakers per speaker type is still a modest amount on the basis of which to speculate on the behavior of the population at large, it is still a sizable improvement, especially given the fact that the participants were well controlled for age, background, etc. (see Section 2.1 below).

Finally and uniquely, the current study expands current research perspectives on pairwise indices and IMs by including corpus-based measures of word frequencies in the analysis. To our knowledge, frequency effects have not been taken into account in the behavior of these metrics, although in light of current research, it seems plausible, and even likely, that durational units are affected by word frequencies. For example, Bell et al. (2009) and Raymond & Brown (to appear) have both shown that other areas of pronunciation are affected by corpus frequencies. Thus, it would be hasty to assume that vowel durations are not

© Servicio de Publicaciones. Universidad de Murcia. All rights reserved.

subject to similar effects. Lemma frequency and word frequency from relevant files of the Corpus del Español (Davies, 2002) were included as independent variables in the statistical analysis of the metrics. Section discusses our methodology, how we obtained syllable duration data from what kinds of speakers etc. as well as the statistical methods we used. Section presents the results of the statistical analysis of the data. Section discusses the significant main effects and interactions in detail, and Section concludes and points to the next steps that should be undertaken.

2. METHODS

2.1. Data

The current study examines and analyzes unscripted speech. Many linguists examine recordings of subjects reading written sentences aloud. This method allows for close control of the material and complications, e.g., the presence of diphthongs or vowel-less syllables, can be controlled for. Furthermore, subjects can be recorded at similar distances from the microphone and can be asked to repeat sentences where pauses or self-correction occurs. However, Deterding (2001) and Carter's various studies of speech rhythm differ in this count. In order to capture more natural speech rhythms, they analyzed recorded spontaneous speech rather than sentences read out loud. Although both spontaneous and scripted speech are still being studied, all other things being equal, natural recorded speech is more likely to reflect spoken language rhythms since "[i]t is well known that there are differences between read and unscripted speech" (Deterding 2001:220). In order to achieve this and following Carter and Detering, we compiled a small specialized corpus consisting of subjects' narrative responses (with minimum interruption on the part of the interviewer) to interviewer prompts.

In order to examine the rhythmic effects of languages in contact in California, the current study used two test groups: ten monolingual Spanish speakers (SPEAKERTYPE: *monolingual*), and ten bilingual English/Spanish speakers (SPEAKERTYPE: *bilingual*). In order to control extraneous variables as much as possible, the study selected test subjects according to specific guidelines: The subjects were between the ages of 18 and 30 and currently enrolled in a four-year university assuring test subjects of a similar age and education level. For both groups, five women and five men were recorded.

The ten monolingual Spanish speakers were intended to represented speakers of a syllable-timed language. To minimize the impact of dialectal variation, speakers in this group were all born and raised in a single region of Mexico, in this case Mexico City, and had never resided in a foreign country. At the time of the study, they were enrolled in a four-year Mexican university, the Universidad Nacional Autónoma de México.

By contrast, the ten bilingual English/Spanish speakers allowed us to measure the effect of bilingualism (with English as the second language) upon the speech rhythm of Spanish in California. The speakers were second-generation Spanish speakers intended to be

© Servicio de Publicaciones. Universidad de Murcia. All rights reserved.

representative of the Chicano population in California. Thus, they were born in California to parents who emigrated from Mexico to California during or after their teen years. Second-generation Spanish speakers generally speak Spanish in the home but learn English outside the home, normally through school. This makes them ideal case studies for bilingualism because they speak both languages from a young age although they are often dominant in English (cf. Montrul 2004a, b, 2005). At the time of the study, speakers were enrolled in upper-division Spanish courses at the University of California, Santa Barbara, which means they either took, or tested out of, two years of Spanish instruction at university level, plus all their coursework and interaction with instructors is in Spanish.

As mentioned above, each group consisted of ten speakers, five males and five females, who were individually recorded for about ten minutes each speaking conversational Spanish, in order to record vowel durations for each speaker. In order to elicit approximately ten minutes of spontaneous narrative speech with a minimum of interruption, each speaker was given a series of prompts chosen to elicit a narrative response. Starting at the beginning of the recording, sentences or phrases displaying a minimal interruption or interference were selected and examined in order to record a minimum of 50 vowel durations per speaker. Vowel duration was recorded for each phrase according to accepted methodology, as described by Mattys & White (2001) for determining the onset and offset of the vocalic nucleus. Accordingly, a visual inspection of speech waveforms and wideband spectrograms using PRAAT phonetic software (Boersma & Weenink 2010) was examined in order to determine and mark the onset and offset of vowels and measure their durations. Given the spontaneous nature of the speech, unnatural syllable elongations due to speaker confusion were eliminated. Regarding Spanish diphthongs, we adopted the methodology employed by Carter (2007). Specifically, Spanish diphthongs were considered as a single vowel. In instances of specific individual complications, such as syllable deletion, these were addressed on a case-by-case basis.

2.2. Statistical Evaluation

To prepare the data for statistical analysis, each syllable in the data was annotated for a number of variables. The dependent variable is SPEAKERTYPE, a categorical variable with two levels, *monolingual* vs. *bilingual*; the following is the list of independent variables:

- SPEAKERSEX: a categorical variable with two levels, *male* vs. *female*;
- IU: a numeric variable ranging from 1 to *n*, where *n* is the number of IUs (intonation units; cf. Du Bois 1991) per speaker; this is included so that we could rule out within-speaker changes over the course of the interview;
- DURATION: a numeric variable providing the length of the vowel in ms;
- SYLLABLE: a numeric variable representing the position of the syllable in the IU; this is only included as a control covariate to make sure that changes over the course of an IU would be controlled for;

© Servicio de Publicaciones. Universidad de Murcia. All rights reserved.

- TOKENFREQ: the log of the frequency of the word form in which the vowel occurred in the Corpus del Español;
- LEMMAFREQ: the log of the frequency of the lemma in which the vowel occurred in the Corpus del Español;
- PVI: the PVI of the duration of the current and the next syllable within the IU (if there was one), computed as in ;
- SD and SDLOG: the standard deviation of the duration of the current and the next vowel within the IU (if there was one) and its natural log (after addition of 1 to cope with 0s);
- VARCOEFF and VARCOEFFLOG: the variation coefficient of the duration of the current and the next vowel within the IU (if there was one), as computed in and its natural log (after addition of 1 to cope with 0s).

(1) $$PVI = \frac{\left|\left(vowel\ duration\ syllable_1 - vowel\ duration\ syllable_2\right)\right|}{mean\left(vowel\ duration\ syllable_{1,} vowel\ duration\ syllable_2\right)}$$

(2) $$VarCoeff = \frac{sd\left(vowel\ duration\ syllable_{1,} vowel\ duration\ syllable_2\right)}{mean\left(vowel\ duration\ syllable_{1,} vowel\ duration\ syllable_2\right)}$$

One comment is necessary regarding the issue of whether/how to control for speech rate. We followed the methodology of Carter (2005, 2007) and Thomas & Carter (2006), whose participants were most similar to those of the current study, and used the raw PVI (*r*PVI) rather than the rate-normalized PVI (*n*PVI). However, the inclusion of the independent variable DURATION nevertheless allows us to explore the statistical relationship between all measures of durational variability on the one hand and speech-rate on the other. This is because DURATION is a proxy for speech rate, with which it is strongly negatively correlated. Thus, any significant interaction between DURATION and any metric would reflect the possibility that a measure's effect is dependent on speech rate. Also, we included logs of durational variability measures to address the possibility of non-linear relationships, something which previous studies have not explored.

To determine how well these variables and their interactions distinguish between the monolingal and the bilingual speakers, all 1061 complete data points were entered into an automatic stepwise bidirectional logistic regression model selection process, trying to predict SPEAKERTYPE: *monolingual*. Using the `stepAIC` function of the R package `MASS` (cf. Ripley 2011 and R Development Core Team 2011), predictors – variables and interactions between them – were added or subtracted until a optimal model was reached, optimal in the sense that it did not benefit from the addition or subtraction of another predictor.

© Servicio de Publicaciones. Universidad de Murcia. All rights reserved.

3. RESULTS

3.1. Overall results and main effects

As a result of the model selection process, several predictors were omitted because they did not contribute enough classificatory power to the model (e.g., IU and TOKEN FREQ). The overall fit of the final regression model to the data is significant (log-likelihood=150.22; df=12; p<0.001), but the classification accuracy is only intermediately good (C=0.704; R^2=0.176; classification accuracy=64.7%); Table 2 provides the coefficients of the final model.

Predictor	Coefficient	*p*	Predictor	Coefficient	*p*
DURATION	-0.02	0.046	DURATION : SYLLABLE	≈-0.001	0.036
SYLLABLE	0.13	<0.001	DURATION : SDLOG	0.006	0.005
SD	-0.04	<0.001	PVI : LEMMAFREQ	0.56	0.001
VARCOEFFLOG	0.47	<0.001	SDLOG : LEMMAFREQ	-0.12	0.017

Table 2: Significant predictors in the final logistic regression model

In what follows, we will discuss these effects (with the exception of SYLLABLE; cf. above). Figure 1 shows the main effects of SD and VARCOEFFLOG on the predicted probability of *monolingual*: As the variability of two vowels increases in terms of SD, the prediction is becoming more likely to be *bilingual*. However, as the variability of two vowel increases in terms of VARCOEFFLOG – i.e., the measure of dispersion less affected by the mean duration – the prediction is becoming more likely to be *monolingual*, at least on the whole (cf. below).

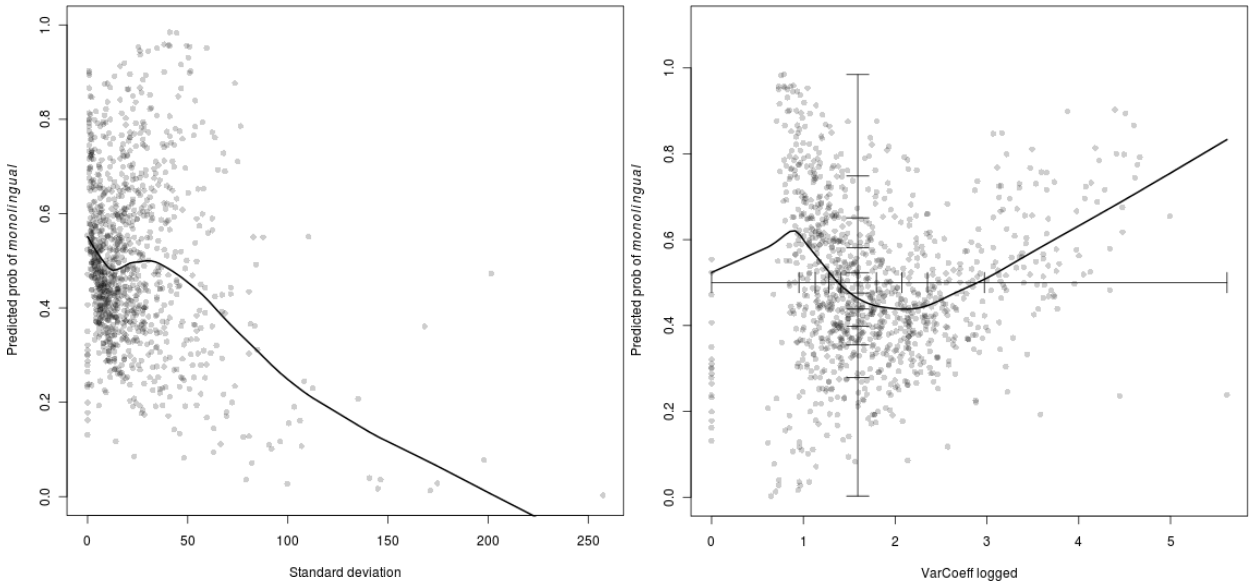

Figure 1: The main effects SD (left) and VARCOEFFLOG (right)
Note: the tick-marked cross in the right panel indicates quantiles

© Servicio de Publicaciones. Universidad de Murcia. All rights reserved.

3.2. Interaction effects

Let us now turn to variables participating in interactions, first, the effect of DURATION. In general, the model reveals that larger durations are ever so slightly associated with bilingual speakers whereas shorter ones are associated with monolingual speakers. However, the significant interactions of DURATION reveal that this variable's effect is much more complex.

Figure 2 represents the interaction DURATION : SYLLABLE: SYLLABLE is represented on the *x*-axis, DURATION is represented on the *y*-axis, each dot represents one observed data point (with grey-shading indicating overplotting), and a linear regression line with a confidence interval summarizes the overall trends for monolingual speakers and bilinguals in the left and right panel respectively.

While Table 2 featured a significant main effect of DURATION (the larger DURATION, the more strongly the model predicts *bilingual*), Figure 2 shows that the two speaker types behave differently. One way of looking at this interaction is that, for monolingual speakers, there is hardly any relation between DURATION and SYLLABLE whereas for bilingual speakers, the later the syllable in the utterance, the longer it tends to be (on average). Another way of putting this is to say that the durations of earlier vowels do not distinguish well between monolingual and bilingual speakers, but that vowels later in the utterance do, because with higher values of SYLLABLE, bilinguals exhibit a trend to have longer durations. In other words, monolingual speakers' vowels late in the utterance are shorter than bilinguals' late vowels. As for "later in the utterance", a non-parametric smoother shows that this interactions begins to manifest itself after 20+ syllables.

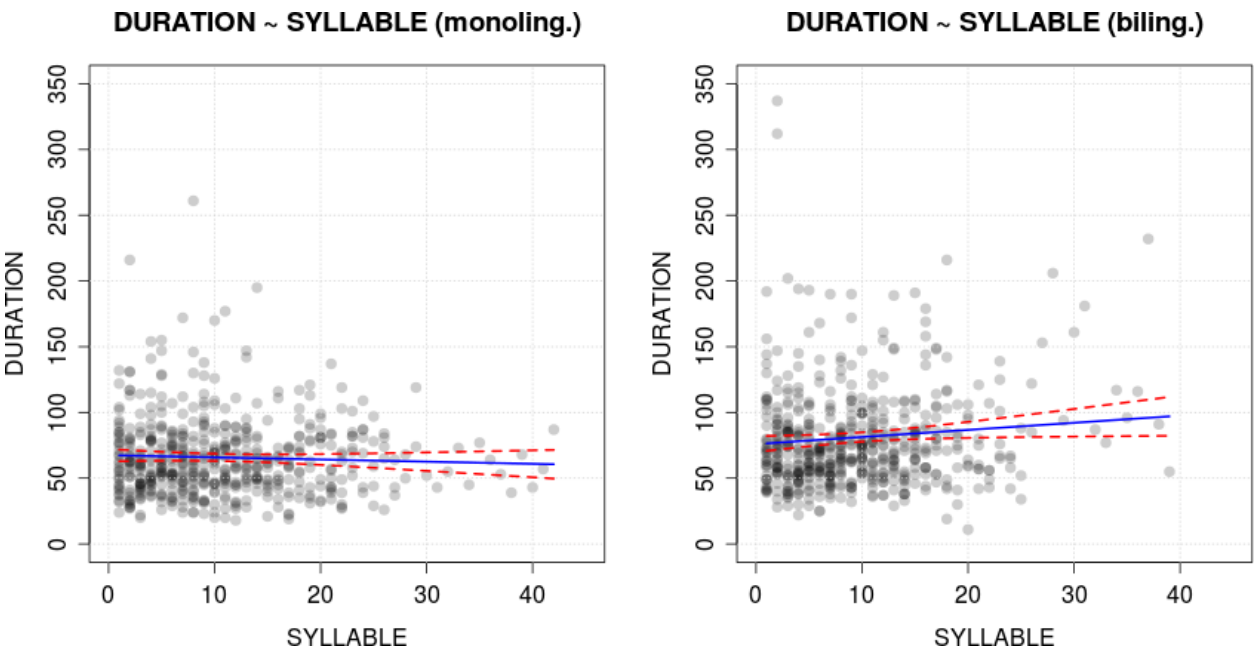

Figure 2: The interaction DURATION : SYLLABLE

© Servicio de Publicaciones. Universidad de Murcia. All rights reserved.

Figure 3 represents the interaction DURATION : SDLOG in a similar format but with an additional twist. Exploration of the data revealed that, for both speaker types, the nature of the correlations between DURATION and SDLOG depends on DURATION as such. For monolingual speakers, as long as syllables are shorter than 72.5 ms, longer durations go hand in hand with reduced variability, but as soon as syllables are longer than that, longer durations go hand in hand with strongly increased variability. For bilingual speakers, syllables shorter than approx. 97.5 ms exhibit are equally variable, after that their variability increased with their length, but not as much as for monolinguals.

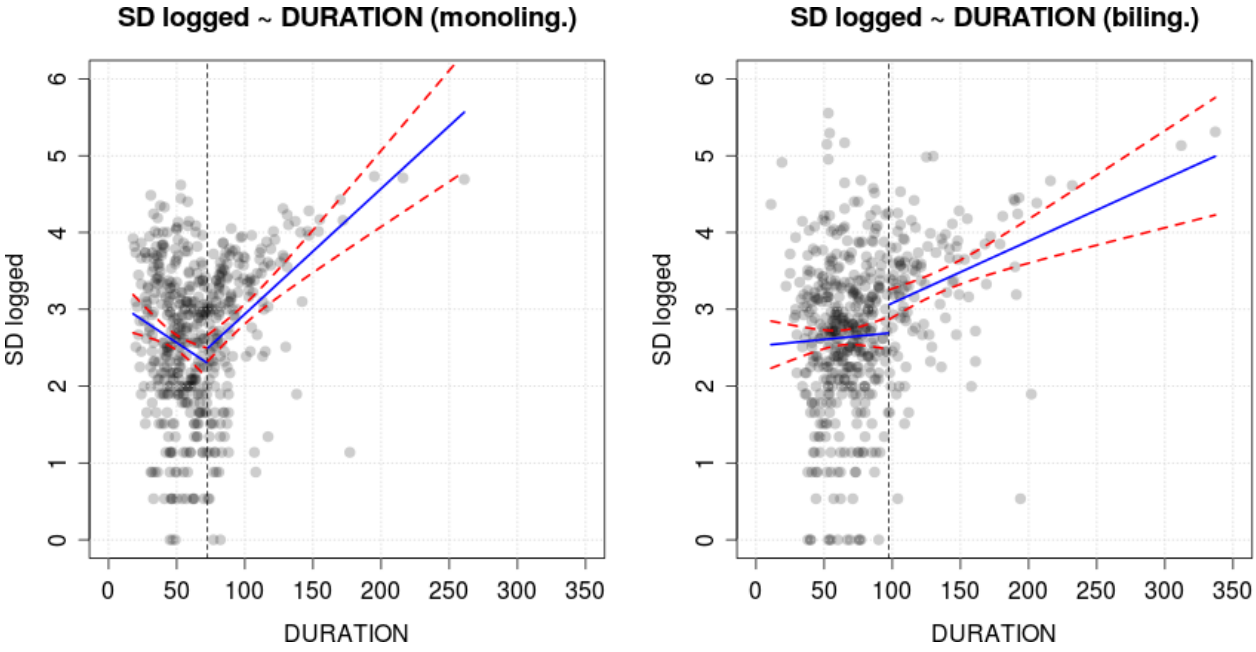

Figure 3: The interaction DURATION : SDLOG

Finally, and interestingly, there are two interactions that involve the corpus-based frequency of the lemma and two ways of measuring the variability of the syllable, the first of which is represented in Figure 4. This shows that the correlation of LEMMAFREQ and SDLOG differs between speakers. More specifically, with high-frequency lemmas, the variability values of mono- and bilingual speakers do not differ, which means SDLOG cannot distinguish the speaker types. However, with words whose lemma frequency is below 9, monolingual speakers have lower SDLOG values.

Finally, Figure 5 represents the interaction LEMMAFREQ : PVI. With medium and high-frequency lemmas, the variability values of mono- and bilingual speakers do not differ, but otherwise the overall trends differ. For monolingual speakers, variability as measured by PVIs is positively correlated with LEMMAFREQ: more frequent words have higher PVIs than less frequent words, but it is the other way round for bilingual speakers. Also, the data show that PVIs can only distinguish mono- and bilingual speakers for words from the extremes of the frequency spectrum: lemmas with LEMMAFREQ<4 and with LEMMAFREQ>9.

In the following section, we will discuss these results in more detail.

© Servicio de Publicaciones. Universidad de Murcia. All rights reserved.

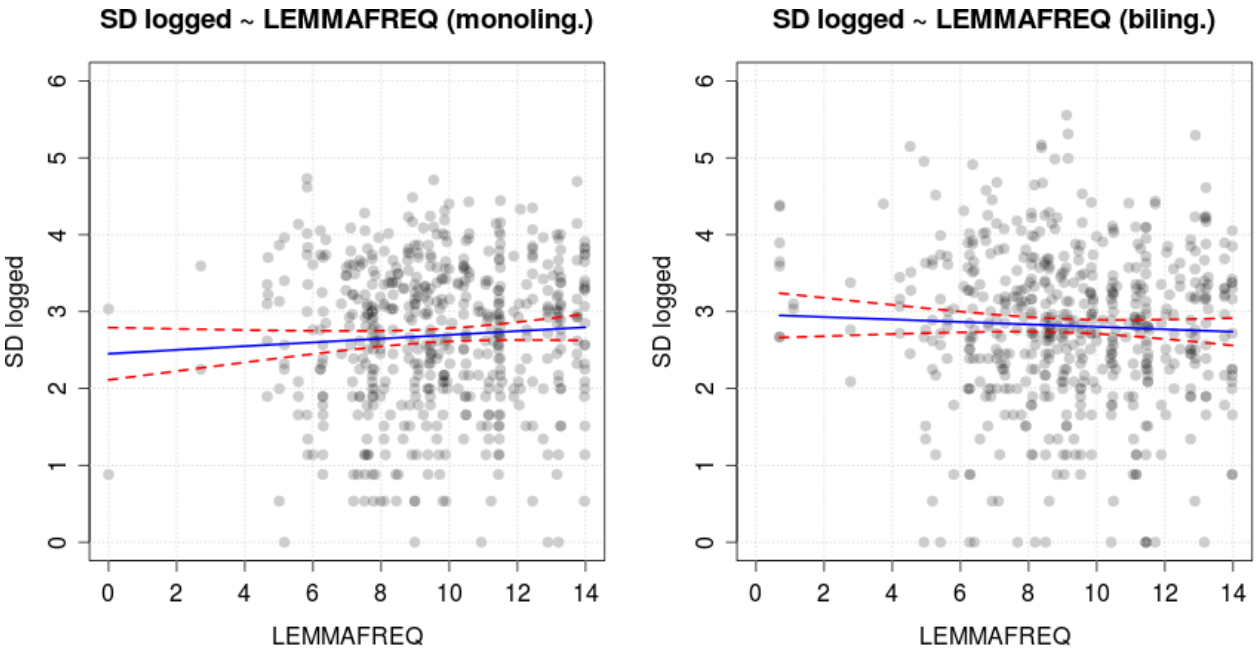

Figure 4: The interaction LEMMAFREQ : SDLOG

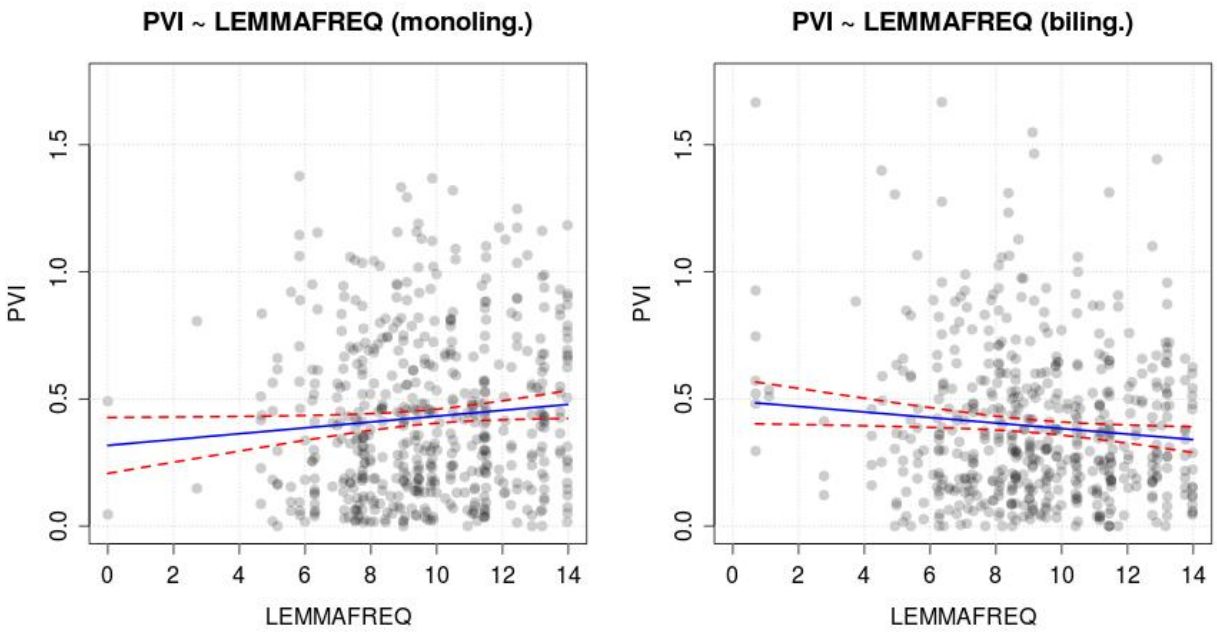

Figure 5: The interaction LEMMAFREQ : PVI

4. DISCUSSION

In this section, first main effects will be discussed, followed by the interactions.

© Servicio de Publicaciones. Universidad de Murcia. All rights reserved.

4.1. Main effects: SD and VarCoeffLog

As mentioned above, both significant main effects, SD and VARCOEFFLOG, are measures of durational variability which *seem to* (cf. below for the reason for the hedging) predict opposite overall trends in speaker type: SD is positively correlated with bilinguality whereas the overall trend of VARCOEFFLOG is negatively correlated with it. Both SD and VARCOEFFLOG are calculated with the standard deviation of the vowel duration within an IU but only the latter controls for the mean syllable duration. We will first discuss the main effect of SD and then VARCOEFFLOG.

It is SD that behaves more as expected: due to the hypothesized influence of English, the bilingual speakers' speech should be more variable in vowel duration than the monolingual speakers' speech, which is exactly what SD reflects. This is compatible with Low & Grabe (1995, 2000), Fought (2003), and Carter (2005, 2007). The former studied L1 vs L2 speakers, finding that Singapore English tended to be more syllable-timed than British English. The speakers in the two latter studies were more similar to those of the current study, in that participants were bilingual and Chicano speakers of Spanish and English, although both studies examined English rather than Spanish. In those studies, the English of Spanish-English bilinguals was more uniform and syllable-timed (i.e. more 'Spanish like') than that of European- and African-Americans, once again suggesting the results reflected by SD, that is, that bilingual speakers would have more variability in vowel duration, reflecting a more 'English like' Spanish. It is important to note that all of the aforementioned studies used the PVI as a metric of durational variability whereas, in the current study, the PVI was not a significant predictor of speaker type (although it did participate in one significant interaction); instead it was SD that reflected the expected influence of bilingualism.

On the other hand, these findings are somewhat at odds with White & Mattys' (2007) conclusions. Upon first glance, SD seems a convenient predictor of vowel durational variability: it is a significant main effect and reflects the expected trend. Its prediction of high variability for bilingual speakers and low variability for monolingual speakers is also compatible with the trend indicated by White and Mattys, who found that Spanish speakers of English showed less vowel durational variability than L1 English speakers (2007). However, this result was indicated by the speech rate normalized standard deviation of vowel duration, rather than the raw standard deviation of vowel duration as in SD. In fact, White & Mattys' (2007:513) found that the standard deviation of vowel duration was not significantly affected by the speakers' native language:

> The results for ΔV [our SD] did not accord with expectations. Both Spanish speakers of English and English speakers of Spanish appeared to realize vocalic interval variation like native English speakers. This does not concur with the subjective perception of these speakers' abilities in the second language: the speakers for both L2 groups were competent but with obvious non-native accents. L1 speakers tended to have faster speech rate than L2 speakers in each case: as seen for the analysis of first languages, there were significant negative correlations between speech rate and ΔV which encourage the use of the rate-normalised VarcoV. [our VARCOEFF].

© Servicio de Publicaciones. Universidad de Murcia. All rights reserved.

However, there are several points that explain and ameliorate the apparent lack of cohesion in the performance of IMs. Firstly, White and Mattys examine L1 vs. L2 speakers, not bilingual speakers. Secondly, they analyze speakers reading a passage, while the current study analyzes more natural, prompted, but otherwise unscripted speech. Thirdly, while the SD differs from VARCOEFF in that it fails to adjust for speech rate, White and Mattys do not include DURATION as a variable in their analysis. As previously mentioned, the inclusion of DURATION approximates speech rate in that higher durations indicate slower speech and vice versa. While SD does not have a speech rate adjustment included in the metric, SDLOG does participate in a significant interaction with DURATION (see Section 3.2). Admittedly, SDLOG is not identical to SD but it is logarithmically derived from it and, hence, perfectly deterministically related. Finally, White & Mattys (2007) do not include frequency effects in their analysis of IMs. LEMMAFREQ participates in two significant interactions with IMs (see Section 3.2), showing that durational variability is clearly effected by corpus-based frequencies. Importantly, one of these interactions is with SDLOG. This suggests that by not including frequency effects in their analysis of the behavior of SD (their ΔV), White & Mattys (2007) cannot provide a complete analysis of the performance of the metrics, making it impossible to determine the true effectiveness of IMs in the context of their study.

Let us now turn our attention to the main effect of VARCOEFFLOG. As mentioned above, at first glance, the trend of VARCOEFFLOG appears to be the opposite of the expected trend and that reflected by SD: Figure 1 suggests an overall positive correlation, according to which monolingual speakers display more variability in vowel duration than bilingual speakers. In other words, it suggests that the Spanish of monolingual speakers is closer in rhythm to English than that of bilingual Spanish-English speakers. However, the overall picture is more complex than a brief glance at the smoother might suggest. Firstly, the smoother in Figure 1 showcases very clearly the danger that linear/straight regression lines come with when forced onto a data set with potential curvilinear trends. In the present case of VARCOEFFLOG, it becomes obvious that, while there is an overall positive correlation – which a linear regression line would have identified, too – this is a case where the prediction is most strongly 'bilingual' in the small range of exactly intermediate variability. Meanwhile, the extreme ranges of variability largely lead to the prediction of 'monolingual'. In fact, as the course of the smoother indicates when related to the quantiles, bilingual speakers tend to group around the mean of VARCOEFFLOG. This would indicate that monolingual speakers are able to employ a full range of vowel durational variability, ranging from zero variability to the most variable syllable pairs, whereas bilingual speakers tend to display an intermediate level of variability according to VARCOEFFLOG, displaying syllable pairs that are neither very similar nor very different.

The fact that native Spanish speakers show a wider range of vowel duration variability than their bilingual counterparts may be related to the monolingual speakers' greater command of the Spanish language. Their aptitude in the use of the language as well as the ability to employ language across a variety of registers may allow them to employ different levels of variability in rhythms in different contexts, leading to the aforementioned effects of VARCOEFFLOG. In comparison, the bilingual Spanish speakers indicated that they primarily

© Servicio de Publicaciones. Universidad de Murcia. All rights reserved.

used Spanish in family situations, so their range of abilities, and perhaps range of durational variability, are likely to be far more constricted. Again, however, this shows that only a more refined type of analysis can identify such elusive patterns.

4.2. The interaction Duration: Syllable

The interaction between DURATION and SYLLABLE is interesting in how it is related to phrase-final syllable lengthening. Speakers tend to elongate syllables towards the end of a phrase with a particularly strong lengthening on the very last syllable of an utterance, a phenomenon generally referred to as prepausal lengthening (cf., e.g., O'Shaughnessy 1995 or Carter 2005). It is precisely because of this type of effect that Low, Grabe, & Nolan (2000) chose to exclude the phrase-final vowel duration in each IU when calculating the mean PVI. In this study, we chose to include all durations but control for this prepausal lengthening by including SYLLABLE as a covariate, which not only avoids the ultimately arbitrary decision as to how many prepausal syllables to exclude (one? two? more?), a question that is particularly pressing as prepausal syllable lengthening appears to be regressive and gradient, but also allows us to explore the data in a more appropriate multifactorial manner with all potential interactions.

Let us first discuss this interaction as it relates to phrase-final lengthening: Phrase-final lengthening leads to the expectation that, as SYLLABLE rises, so would DURATION. However, while bilinguals do exhibit this expected trend, the Mexican monolingual speakers actually display the opposite one, with shorter vowels as SYLLABLE rises.

One possible explanation for the seemingly unexpected trend of the monolingual speakers is related to Mexican vowel reduction. First and in general, certain dialects of Mexican Spanish have been shown to have marked vowel reduction. While Lope Blanch's (1963) seminal investigation of Mexican vowel reduction does not show a consistent vowel reduction amongst Mexican speakers, it does confirm the presence of at least intermittent vowel reduction amongst a majority of participants (Lope Blanch 1972:7). In our data, the average duration of monolingual speakers' vowels is shorter than that of bilingual speakers (the significant effect of DURATION, p=0.046).

Second and with regard to this interaction in particular, other studies have yielded similar results. For example, Delforge (2008:115) stated that Spanish unstressed vowel reduction is gradient in effect, variable in occurrence, and associated with word final syllables. The author further states that, in Andean Spanish, for example, vowel reduction in prepausal syllables is particularly severe, with 87% of vowels being either completely devoiced or apparently deleted. While our data are on Mexican, not Andean Spanish, our interaction DURATION : SYLLABLE is perfectly compatible with her earlier results. This indication of the presence of Mexican vowel reduction in particular positions in the IU suggests that further research along these lines may be beneficial in our understanding of prosodic patterns of bilingual and monolingual Spanish.

An additional or alternate explanation for the tendency of bilingual speakers to display longer vowel durations as compared to monolingual speakers, and especially late in the

© Servicio de Publicaciones. Universidad de Murcia. All rights reserved.

utterance, may be that the lesser degree of proficiency of the bilingual speakers results in a larger amount of processing cost, especially during the online processing of linguistic material later in an utterance, when content and grammatical constraints (e.g., agreement) make it harder to continue and complete a structure begun much earlier. At this point, however, this is merely speculative.

4.3. The interaction Duration: SDLog

The interaction DURATION: SDLOG revealed that, on the whole, as DURATION increases, SDLOG increases. This overall trend is compatible with White & Mattys (2007), who found a negative correlation between SD and speech rate. In our case, our approximate of speech rate, DURATION, is negatively correlated with speech rate, thus high variability according to SDLOG is associated with high values of DURATION. However, the interaction was somewhat more complex. Two main findings need to be discussed. First, the interaction revealed that monolingual speakers exhibit higher SDLOG values for longer syllables than bilinguals. This is apparent from the right regression line on the left panel of Figure 3, which reveals a much steeper trend than the right one on the right panel, indicating increasingly higher SDLOG values for monolingual speakers as the length of the syllable increases.

Second, bilingual speakers' short vowels are all equally variable, but monolingual speakers' ones are not. This emerges from a comparison of the left regression lines in both panels: For monolingual speakers' short vowels (i.e., 72.5 ms or less), as they become longer, they become less variable (the regression line goes down). For bilingual speakers' short vowels (i.e., 97.5 ms or less), the vowels do not change in terms of variability.

What explains the disparate trends in the variation of short vowels by bilingual and monolingual Spanish speakers? While this is speculative, these data may be another example of how differences in linguistic aptitudes are reflected between speaker types. Bilingual speakers may produce pronunciation that is slower and more careful but also more homogeneous (when controlled for DURATION): When the vowels are short they do not vary much and they do not change as their length changes, and when vowels become longer, these speakers do not exhaust the full variability range as much as monolingual speakers do. Monolingual speakers, on the other hand, appear to exhibit a minimal level of variability at one particular (typical?) syllable duration, but can make use of increasing or decreasing syllable lengths and variability (for emphatic purposes) more flexibly than the bilingual speakers.

One implication of the variable behavior of IMs across different vowel durations reveals that simply reporting a measure of central tendency for metrics of durational variability is risky in that it implies that these metrics behave more or less identically across different levels of speech rate. While some metrics do include rate-normalization coefficients, the exploration of interactions such as DURATION : SDLOG allows for a more complete picture of the behavior of vowel durations across different speech.

© Servicio de Publicaciones. Universidad de Murcia. All rights reserved.

4.4. The interaction LemmaFreq: SDLog

The two interactions involving word frequency both prove to be highly interesting as well as important in their implications for further research of speech rhythms. The first, LEMMAFREQ: SDLOG, indicates that monolingual speakers exhibit less variability in vowel duration (measured in SDLOG) for less frequent words. In other words, with common words, bilingual speakers behave like monolingual ones, but with uncommon words, bilingual speakers are less homogeneous. This may be explained by linguistic aptitude: on average, bilinguals will have less exposure and practice – in terms of both comprehension and production – and, thus, speak more slowly. At the same time, it seems that their lesser proficiency also manifests itself in more heterogeneous production especially for those words to which they are even less exposed to: words of low frequency.

4.5. The interaction LemmaFreq: PVIç

The final interaction involves lemma frequency again, but this time with a different durational variability measure, the PVI. However, SDLOG and PVI themselves are positively (and exponentially) related ($PVI \approx 0.02 * 2.609^{SDLOG}$; $R^2=0.86$), which is why it is not surprising to see that this interaction is similar to LEMMAFREQ : SDLOG. Again, with less frequent words, monolingual speakers' durational variability is lower than that of native speakers. However, the present interaction shows that the PVI's effect is frequency-dependent just like that of SDLOG, but for a different range of lemma frequencies: SDLOG cannot distinguish speaker types with high frequency lemmas, but PVI can, but SDLOG can distinguish speaker types with medium-frequency lemmas, which the PVI cannot.

Since two measures of durational variability interact with LEMMAFREQ, this raises the question of how they compare to each other. On the one hand, it seems as if the PVI can distinguish the two speaker types over as wide a range of lemma frequencies as SDLOG, even if it is two non-consecutive ranges, high- and low-frequencies, but not intermediate ones. However, it must be borne in mind that frequency ranges of words are not all equally populated: frequencies are Zipfian-distributed, which means that there are very many words of low frequency, intermediately many words of medium frequency, but only very few words of high frequencies. Thus, the fact that the PVI can distinguish speaker types for high-frequency lemmas better than SDLOG does not make it a more appealing measure because that will only include very few lemma types – by contrast, the fact that SDLOG can distinguish speaker types for all lemma types with a frequency of less than 9 makes it a more widely applicable measure.

4.6. Implications

The findings discussed above yield several implications. First and on a very general level, the present data indicate that measures of durational variability are related to each other and to speaker types, frequencies, etc. in extraordinarily complex ways involving main effects, interactions, nonlinear effects and breakpoints. Against this background, it is somewhat

© Servicio de Publicaciones. Universidad de Murcia. All rights reserved.

amazing that previous, quantitatively less thorough work lead to the interesting results that it did.

Second, with regard to the final model it is clear that such measures interact with corpus-derived lemma frequencies. And in a sense, this is not really surprising given how corpus frequencies affect other aspects of pronunciation including, but not limited to, speed of articulation or thoroughness of articulation (or its inverse, amount of reduction); cf., e.g., Bell et al. (2009) or Raymond & Brown (to appear). It is interesting to note in this connection, however, that it is *lemma*, not token, frequency that is more relevant to the speakers in the present data, which is surprising since usually word/token frequencies are more decisive for process of articulation and, say, historical change, an issue to which we will return briefly in the following and final section. Regardless of which type of frequency will turn out to be more relevant to durational variability, future studies should not only try to approach durational variability in quantitatively more advanced ways (i.e., multifactorially) but also take frequency effects based on corpus data into consideration.

Third, the data as well as some additional arguments shed some doubt on the utility of the PVI. With regard to the data, the PVI does not feature as a main effect in the final regression model and only features in one interaction (with LEMMAFREQ); its classificatory power is therefore more limited than that of other predictors. Also, even within said interaction, the PVI's classificatory power is restricted to a smaller subset of the data than the competing measure of SDLOG: the range of words for which the PVI will be useful is smaller than that of SDLOG. To these empirical findings, we may add conceptual, or design, weaknesses of the PVI. On the one hand, the PVI as often used is a mean of means of means. However, it is well-known that means are really only appropriate measures of central tendencies for normally-distributed data, and in our data, the PVIs of 18 of the 20 speakers are significantly different from normality (and one of the two remaining speakers' PVIs are very close to that, too, with $p_{\text{Shapiro-Wilk test}}=0.051$). Relatedly, the 'nested averaging' simply ignores a lot of variability that a more comprehensive account would want to account for. For example, the 'nested averaging' also does not allow one to study PVIs on a syllable-by-syllable basis (since only an average will be considered in the traditional approach), but that rules out the incorporation of, for instance, frequency effects for lemmas (as here), words, syllables as well as other lexically-specific predictors.

As is often the case with a multifactorial investigation of a complex data set, the results and their interpretations are quite complex. The current study did reveal a more complicated perspective of the prosodical differences between monolingual and bilingual speakers than those previously reported by Carter (2005, 2007) or the perspectives on L1 and L2 speakers reported by Low, Grabe, & Nolan (2002). However, this is not surprising: While the attempts to quantify rhythmic differences have not utilized quantitatively sophisticated tools, no current work claims rhythmic differences are simple or clear cut – in fact, returning to the exploration of the PVI and IMs, some researchers suggest that vowel duration variability is only a small part of the overall picture of language rhythms (cf. Kohler 2009) and the current study reveals that multiple facets of vowel durational variability must be

© Servicio de Publicaciones. Universidad de Murcia. All rights reserved.

clearly distinguished and integrated if we want to address the larger challenge of quantifying speech rhythms.

However, the current study also yields interesting results regarding the prosodies of two related dialects of Spanish, Californian Chicano Spanish and Mexican Spanish. As mentioned previously, the picture is quite complex. Monolingual Spanish speakers and bilingual Spanish-English speakers displayed differences in vowel durational variability but the dialects did show similarities as well: both displayed similar overall correlations between vowel length and position in the intonational unit, and both showed similar levels of variability at certain word frequencies. Certainly the question as to what extra-dialectal factors influenced these results (specifically speaker fluency levels and, maybe, online processing constraints) remain to be further explored. It is such similarities and their underlying causes the exploration of which will advance future research to help us gain a more complete perspective of language rhythms.

5. WHERE TO GO FROM HERE

Several next steps suggest themselves. On the one hand, there are the obvious ones. With regard to the issue of how language rhythms differ between mono- and bilingual speakers, replications using more kinds of monolingual and bilingual speakers would be desirable, where the languages would ideally also be from opposing ends of the language-rhythm continuum. With regard to the utility of different measures, such data as well as data more diverse in terms of spoken registers could be useful to explore which of the measures result in the largest discriminatory power, as would maybe statistically more refined analyses of existing data on the PVI, SD, VARCOEFF, and maybe others.

In the context of this special issue the role of corpus frequencies is particularly relevant. This study was the first to study the PVI and other measures by including frequency effects. However, in spite of this new methodological advantage, this could only constitute a very first step. It is already well-known that measures more specific than mere frequency of occurrence play a role in other articulatory phenomena, such as transitional probabilities, association measures such as *Mutual Information*, contextual predictability, the difference between content and function words, etc., and it stands to reason that durational variability *per se* could be equally affected by such factors. In addition, even existing work on the above issues could benefit from reanalysis. For example, corpus-based association measures have been shown to be relevant to articulation, but have so far also been virtually exclusively bidirectional, but Gries (under review) shows that directional measures can paint a more accurate picture of lexical association, which may affect durational variability.

More importantly even, there is recent work that strongly supports the notion that simple frequency counts are much less important than the kind of higher-dimensional data made of multidimensional conditional probabilities, entropies, etc.: In a truly pioneering study, Baayen (2010) provides comprehensive evidence for the assumption that the kind simple frequency effects corpus linguists and psycholinguists often invoke merely arise out of

© Servicio de Publicaciones. Universidad de Murcia. All rights reserved.

learning a wide range of more specific distributional patterns in the context of expressions, and given my above pleas, the following is worth quoting in detail (our emphases):

> A principal components analysis of 17 lexical predictors revealed that most of the variance in lexical space is carried by a principal component on which *contextual measures* (syntactic family size, *syntactic entropy*, *BNC dispersion*, morphological family size, and *adjectival relative entropy*) have the highest loadings. Frequency of occurrence, in the sense of pure repetition frequency, explains only a modest proportion of lexical variability. Furthermore, the principal component representing local syntactic and morphological diversity accounted for the majority of the variability in the response latencies … (Baayen, 2010:456; cf. also Raymond & Brown, to appear)

Findings like these indicate why measures such as the PVI may be too simplistic – the multiple averaging decontextualizes all variability – and why even the present approach can only be a starting point to explore durational variability in the rich and authentic contexts in which it occurs and with the whole arsenal of corpus-based metrics and statistical techniques that are available. If this paper stimulates research on durational variability along some such lines and involving such a more corpus-informed perspective, it has achieved one of its main goals.

ACKNOWLEDGEMENTS

We thank UC MEXUS for a research grant made available to the first-named author as well as Viola G. Miglio for valuable feedback and comments. The usual disclaimers apply.

REFERENCES

Abercrombie, D. (1967). *Elements of general phonetics*. Edinburgh: Edinburgh University Press.

Baayen, R. H. (2010). Demythologizing the word frequency effect: a discriminative learning perspective. *The Mental Lexicon, 5*(3), 436-461.

Bell, A., Brenier, J. M., Gregory, M., Girand, C., & Jurafsky, D. (2009). Predictability effects on dirations of content and function words in conversational English. *Journal of Memory and Language, 60*(1), 92–111.

Boersma, P. & Weenink, D. (2010). Praat: doing phonetics by computer (Version 5.1.29 [Computer program]. Retrieved June 21, 2009, from <http://www.praat.org/>.

Bunta F. & Ingram, D. (2007). The acquisition of speech rhythm by bilingual Spanish- and English-speaking 4- and 5-year-old children. *Journal of Speech, Language, and Hearing Research, 50*(4), 999-1014.

Carter, P. M. (2005). Quantifying rhythmic differences between Spanish, English, and Hispanic English. *Theoretical and Experimental Approaches to Romance Linguistics: Selected Papers from the 34th Linguistic Symposium on Romance Languages* (pp. 63–75). Amsterdam: John Benjamins.

© Servicio de Publicaciones. Universidad de Murcia. All rights reserved.

Carter, P. M. (2007). Phonetic variation and speaker agency: Mexicana identity in a North Carolina Middle School. *University of Pennsylvania Working Papers in Linguistics 13*(2), 1–14.

Dasher, R. & L. Bolinger, D. L. (1982). On pre-accentual lengthening. *Journal of the International Phonetic Association,* 12, 58–69.

Dauer, R. M. (1987). Phonetic and phonological components of language rhythm. *Proceedings of the 11th International Congress of Phonetic Sciences*, Vol. 5, pp. 447–450. Tallinn.

Davies, M. (2002-). Corpus del Español: 100 million words, 1200s-1900s. <http://www.corpusdelespanol.org>.

Delforge, A. M. (2008). Unstressed vowel reduction in Andean Spanish. In L. Colantoni & J. Steele (Eds.), S*elected proceedings of the 3rd Conference on Laboratory Approaches to Spanish Phonology* (pp. 107–124). Somerville, MA: Cascadilla Proceedings Project.

Deterding, D. (2001). The measurement of rhythm: a comparison of Singapore English and British English. *Journal of Phonetics*, *29*(2), 217–230.

Du Bois, J. W. 1991. Transcription design principles for spoken discourse research. *Pragmatics 1*(1), 71–106.

Fought, C. (2003). *Chicano English in context*. London: Anthony Rowe Ltd.

Gries, S. Th. (under review). 50-something years of work on collocations: what is or should be next … *International Journal of Corpus Linguistics*.

Kohler, K. (2009). Whither speech rhythm research? *Phonetica, 66*(1–2). 5–14.

Lope Blanch, J. M. (1972). *Estudios sobre el español de México*. Universidad Nacional Aútonoma de México. México D.F., México.

Low, E., Grabe, E., & Nolan, F. (2000). Quantitative characterizations of speech rhythm: syllable-timing in Singapore English. *Language and Speech*, *43*(4), 377–401.

MacLeod, A.A.N. & Stoel-Gammon, C. (2005). Voice onset time (VOT) in Canadian French and English: Monolingual and bilingual adults. *Journal of the Acoustical Society of America, 117*(4), 2429-2429

Montrul, S. (2004a). Convergent outcomes in second language acquisition and first language loss. In M. Schmid, B. Köpke, M. Keijzer, & L. Weilemar (Eds.), *First language attrition* (pp. 259–280). Amsterdam & Philadelphia: John Benjamins.

Montrul, S. (2004b). Subject and object expression in Spanish heritage speakers: A case of morphosyntactic convergence. *Bilingualism: Language and Cognition*, *7*(2), 125–142.

Montrul, S. (2005). Second language acquisition and first language loss in adult early bilinguals: Exploring some differences and similarities. *Second Language Research, 21*(3), 199–249.

Sorace, A. (2004). Native language attrition and developmental instability at the syntax-discourse interface: Data, interpretations and methods. *Bilingualism: Language and Cognition, 7*(2), 143–145.

O'Shaughnessy, D. (1995). Timing patterns in fluent and disfluent spontaneous speech. *Acoustics, Speech, and Signal Processing*, *95*(1), 600–603.

© Servicio de Publicaciones. Universidad de Murcia. All rights reserved.

Pike, K. (1945). *The intonation of American English*. Ann Arbor: University of Michigan Press.

R Development Core Team. (2011). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. <http://www.R-project.org/>.

Ramus, F. (2002). Acoustic correlates of linguistic rhythm: Perspectives. In *Proceedings of speech prosody 2002*, 115–120. Aix-en-Provence.

Ramus, F., Nespor, M., & Mehler, J. (1999). Correlates of linguistic rhythm in the speech signal. *Cognition, 73*(3), 265–292.

Raymond, W. D. & Brown, E. L. (to appear). Are effects of word frequency effects of context of use? An analysis of initial fricative reduction in Spanish. In Stefan Th. Gries & Dagmar S. Divjak (Eds.), *Frequency effects in language: learning and processing*. Berlin & New York: Mouton de Gruyter.

Ripley, B. (2011). *MASS*. Version 7.3-13 Package for R.

Thomas, E. & Carter, P. M. (2006). Prosodic rhythm and African American English. *English World Wide, 27*(3), 331–355.

White, L. & Mattys, S. L. (2007). Calibrating rhythm: First language and second language studies. *Journal of Phonetics*, 35(4), 501–522.

© Servicio de Publicaciones. Universidad de Murcia. All rights reserved.