What does %V actually measure?

Margaret E. L. Renwick[*]

1 Introduction

The system of classifying languages by their rhythmic patterning traditionally and historically characterizes the patterns in terms of three groups. One group includes the so-called "stress-timed" languages like English, German, and Dutch, which are perceived as having equal intervals between stresses, and are described as sounding rhythmically like Morse code (James 1940). A second rhythmic class includes the "syllable-timed" languages such as Spanish, Italian, and French, described by James (1940) as having a staccato beat, in which syllables are perceived as coming along at regular intervals. Abercrombie (1965) assigned Russian and Arabic to the former category, and Yoruba and Telugu to the latter. Japanese is argued to belong to the third, "mora-timed" class, as investigated phonetically by Port et al. (1987), among others. Several languages, such as Catalan and Polish (Dauer 1987; Nespor 1990) have been claimed to be intermediate languages, exhibiting phonological characteristics of both the stress- and syllable-timed classes, and many languages remain unclassified.

While these traditional rhythmic classifications are tied to ideas of isochrony, i.e. the temporally regular occurrence either of stresses or of syllables, studies intended to confirm these intuitions by tying rhythmic properties directly to speech patterns, i.e. in terms of measurable intervals between stresses or syllables, have been inconclusive at best (Lehiste 1973, 1977; Roach 1982). Thus, although the perception of regular speech rhythm is often described, it has been exceedingly difficult to tie perceptions of rhythm to quantitative periodicity in any one particular phonetic aspect of a language (i.e. duration); instead, it has been hypothesized that rhythmic differences may at best be captured by a combination of factors, such as possible syllable shapes and the presence or absence of vowel reduction (Dauer 1983).

Nevertheless, in recent decades the long-standing intuition that languages differ in their speech rhythm, from staccato 'machine-gun-like' strings of rhythmically equivalent syllables to alternating patterns of strong and weak syllables, has received apparent support from both perceptual and acoustic studies. For example, Ramus, Dupoux and Mehler (2003) reported a high success rate for listeners in distinguishing among the rhythm patterns of severely downsampled utterances from languages of different classes, supporting a proposed acoustic measure of the durational ratio of vocalic regions to total acoustic content, %V,[1] which seemed to similarly distinguish among languages of the different proposed rhythmic types. Ramus, Nespor and Mehler (1999), among others, found that languages traditionally described as stress-timed (e.g. Dutch, English, Russian) tended to have lower %V values than syllable-timed languages (e.g. French, Italian, Spanish); and mora-timed Japanese had the highest %V of all.

However, subsequent reports using %V as a measure have differed considerably from the original results (White and Mattys 2007), and Arvaniti (2009) showed that sentences containing many open syllables tend to have a higher %V than those that contain many closed and complex syllables, regardless of language. This paper quantifies that relationship by examining the correlation between %V and syllable structure in a set of spoken utterances whose sentence targets were

[*] Thanks go to Abby Cohn, Stefanie Shattuck-Hufnagel, Michael Wagner and Draga Zec for guidance, helpful suggestions and comments on this paper and its experimental basis. I would also like to thank participants in the 2008 Research Workshop, Cornell University Linguistics Department. Thanks to Johanna Brugman and Bill Renwick for technical support. Any remaining errors are my own. Finally, I thank the speakers who participated in the study described here.

[1] Abbreviations: %V = the ratio of vocalic material to the total duration of an utterance; ΔC = the standard deviation of consonant intervals over an utterance

drawn from the set originally tabulated by Ramus et al. (1999); the relationship is modeled by calculating the number of open vs. closed syllables in a target sentence (where a sentence's syllable count is based on the lexical form), as well as the number of complex onsets.

The methodology behind %V, a calculation which relates the sum total of the durations of vocalic intervals in a spoken sentence to its overall duration, was inspired by perceptual experiments showing that participants can discriminate between sentences of resynthesized speech from different 'rhythm classes', but cannot do so for languages in the same class. These results have been important in recent cognitive science research, as they have been replicated in studies of discrimination by tamarind monkeys (Tincoff et al. 2005) and by rats (Toro et al. 2003), demonstrating that the faculty to perceive speech rhythm as defined in this way is not limited to humans.

Meanwhile, researchers interested in the role of rhythm in language acquisition have also taken an active interest in these issues. In a series of experiments using low-pass filtered speech, Nazzi et al. (1998) showed that infants could distinguish between languages in different traditionally-defined rhythm classes, but not between languages in the same class (like English and Dutch). Ramus and Mehler (1999) and Ramus et al. (2003) extended these findings by showing that adults could discriminate between languages in different classes, even when all intonational information was removed from the speech signal, and strings of vowels and consonants were replaced by a sequence of *sa – sa – sa* syllables. As a phonetic correlate to the durational cues that allowed such discrimination, Ramus et al. (1999) proposed %V, which measures the durational ratio of vocalic to consonantal material in an utterance.

For their original experiment exploring the possibility that the %V measure could capture listener intuitions about the rhythmic differences among languages, Ramus et al. (1999) used four speakers for each of eight languages, and chose five sentences per speaker for analysis. They found that utterances in Dutch and English (designated as stress-timed languages) and Polish (a rhythmically-unclassified language) had low %V values; Spanish, Italian and French (designated as syllable-timed languages) and Catalan (unclassified) had higher values; and Japanese (a mora-timed language) had a much higher %V; they further found that a language's rhythmic classification was a significant predictor of its %V and its ΔC (the standard deviation of consonant intervals over a sentence).

Subsequent studies using the %V methodology have generally found the same ordering of languages with respect to %V: it is high for Japanese, but low for English, and in-between for languages like Spanish. However, the precise %V measurements have not been replicable, and given the relatively small differences in %V between languages, these discrepancies have raised some questions about the usefulness of a measure like %V. This has led other investigators to try various means of making this measure more replicable. For example, Dellwo and Wagner (2003) implemented controls for speech rate (though they found little correlation between speech rate and %V); and Dellwo (2006) took into account variability in consonant interval duration, which helped distinguish between stress-timed English and German, vs. syllable-timed French.

White and Mattys (2007) compared a variety of rhythm measures, including %V, recording their own set of utterances for this study. Their stimulus sentences were constructed with as few approximants as possible, to ease the process of segmenting consonants from vowels; the original study by Ramus et al. (1999) had no controls on segmental content, although in both studies each sentence had roughly the same number of syllables. White and Mattys' %V values for English and Dutch were slightly lower than those of Ramus et al., while French and Spanish both had higher %V values than in the original study. This is likely due to the constraints placed on the segmental context of the experimental stimuli. That is, in Spanish, unlike in English, not all natural classes may appear in consonant clusters, and approximants are among the few that can appear in syllable codas. If words in the stimuli are constrained not to contain approximants, particularly /l/ and /r/, then there will be fewer consonant clusters, a higher instance of open syllables, and by extension a higher %V. In fact, White and Mattys' (2007) Spanish stimuli contain very few closed syllables – a result that already indicates a significant link between phonotactics, specifically syllable shape, and the %V value obtained for a particular sample of spoken utterances.

In a similar vein, Arvaniti (2009) explicitly tested the role of syllable structure in determining the value of %V. She systematically varied the syllable shapes in utterances and then measured their %V. She compared English (stress-timed), Spanish (syllable-timed) and Greek (rhythmically unclassified), using a variety of phonotactically-controlled sentences in each language. The sentences for each target language were constructed to be either "stress-timed," i.e. with as many consonant clusters as possible, so that individual syllables varied widely in their phonotactic shape; "syllable-timed," with a majority of open syllables and minimal clusters, so that individual syllables were more consistent in their phonotactic shape; or uncontrolled. They found that the value obtained for %V was highly dependent on the specific materials recorded: in each language, the "stress-timed" sentences had the lowest %V, the "syllable-timed" sentences had the highest, and the uncontrolled sentences fell in between. Taken together, the results of White and Mattys on the one hand, and Arvaniti on the other, show that sentences containing many open syllables tend to have a higher %V than those trending towards closed or complex syllables, and that this holds regardless of language.

The experiment presented in the next two sections adds to the evidence that the %V value for any individual language is a highly variable number that may not help to quantify the relatively constant percept of cross-linguistic timing differences, unless the corpus is large enough to reflect the statistical distribution of syllable types in that language. In addition, this study indicates what %V *does* do: it correlates with the syllable phonotactics of an utterance, most directly with regard to codas, but also with regard to complex onsets. This experiment is complementary to Arvaniti (2009) in the sense that, instead of designing target sentences to contain particular syllable structures, it selects targets from Ramus et al.'s original set for a number of languages, elicits new utterances of those targets, and then correlates the syllable structure of those utterances with their %V values.

2 %V: New data from five languages

We now turn to data from this new experiment using the %V methodology, to further test the relationship between syllable structure and %V in five languages: English, Dutch, Italian, Spanish and Japanese. The experiment is motivated by the hypothesis that if rhythm class, as traditionally understood, plays a role in determining the value of %V, then %V should be highest for Japanese, slightly lower for syllable-timed Italian and Spanish, and lowest for stress-timed English and Dutch, in any given utterance produced by speakers in those languages.

These %V measurements were carried out on new recordings of a sample of target sentences drawn from Ramus et al.'s (1999) stimuli (originally recorded by Nazzi et al. 1998). These new utterances were collected in the Phonetics Laboratory at Cornell University from two speakers each of English, Dutch and Italian, and three speakers each of Spanish and Japanese. For English, Italian and Spanish, all speakers were female; for Dutch and Japanese, one speaker was male and the others were female. At least 20 sentences were analyzed per language (10 sentences/speaker); the sentences were chosen at random from the original set recorded by Nazzi et al. (1998); the subset of analyzed sentences was *not* identical across speakers of each language (see Appendix B for speaker- and sentence-specific results). These sentences were selected from the set of sentences described by Ramus et al. (1999), but because those authors analyzed only a subset of their recordings and do not specify which subset of utterances was analyzed, our selection almost certainly differs from theirs. Thus, if the ordering of %V values for the five languages is maintained in this study, and if we find magnitudes of difference in %V across languages that are comparable to those in other studies, it will provide support for this measure as a diagnostic for rhythm class. On the other hand, if the ordering is different for these recordings, or if %V relates systematically to the CV structure of individual utterances rather than to the language per se, it will suggest that much of a %V value is determined by that syllable structure. Results of the %V analysis for each language appear in Table 1, and they show that while the original results reported by Ramus et al. are generally replicated, there are some crucial differences here. For comparison, the original %V results found by Ramus et al. (1999: 272), for the languages also reported here, are reproduced in Table 2.

The languages shown in Table 1 and Table 2 are arranged in order of %V value, from lowest to highest, and we see that in general, this replication produced lower %V numbers than the original study. We also note that the difference in %V between English and Spanish is the smallest difference between any pair of adjacent languages in this table; and that the biggest difference in %V between two adjacent languages is between Italian and Spanish. This is surprising, given that the large proposed rhythmic difference between English and Spanish should be reflected in different %V values, while Spanish and Italian should be quite similar. Nevertheless, when (as in the original study) an ANOVA was run to test a model in which Rhythm Class predicted %V, the resulting R^2 was 30.1%, and Rhythm was a significant factor ($F(2) = 27.2133$; $p < .0001$).

Language	Vocalic Intervals	Consonantal Intervals	Number of Sentences	**Mean %V (SD %V)**	Mean ΔV (SD ΔV)	Mean ΔC (SD ΔC)
Dutch	283	290	20	39.606 (4.9)	4.26 (1.05)	6.84 (1.91)
English	362	379	25	41.701 (6.0)	5.57 (1.49)	8.45 (2.35)
Spanish	463	468	30	42.034 (5.6)	3.75 (1.95)	5.43 (1.64)
Italian	422	412	24	47.657 (3.9)	4.40 (0.96)	5.05 (0.88)
Japanese	458	457	30	50.036 (4.3)	4.20 (1.03)	4.44 (1.0)

Table 1. Experiment 1: %V Results by Language[2]

Language	Vocalic Intervals	Consonantal Intervals	Number of Sentences	Mean %V (SD %V)	Mean ΔV (SD ΔV)	Mean ΔC (SD ΔC)
English	307	320	20	**40.1** (5.4)	4.64 (1.25)	5.35 (1.63)
Dutch	320	329	20	**42.3** (4.2)	4.23 (0.93)	5.33 (1.5)
Spanish	320	317	20	**43.8** (4.0)	3.32 (1.0)	4.74 (0.85)
Italian	326	317	20	**45.2** (3.9)	4.00 (1.05)	4.81 (0.89)
Japanese	336	334	20	**53.1** (3.4)	4.02 (0.58)	3.56 (0.74)

Table 2. %V Results of Ramus et al. (1999: 272)

As a hypothesis-independent measure of how the data group languages with proposed rhythmic similarities together, a Tukey test was performed (as had been done in earlier studies) to evaluate the significance of differences in %V across rhythm classes, by comparing the mean %V values of each rhythm class. Results showed that the three "rhythm classes" are significantly different in terms of mean %V. In addition, a second ANOVA tested the effect of Language on %V, which had not been tested in earlier studies. This model, in which the factor Language was significant ($F(4) = 20.1091$, $p < .0001$), had an R^2 of 39.3%, considerably higher than the model using rhythm as its factor, indicating a better fit to the data. That is, when the data are grouped by language, the resulting model accounts for more of the variability in %V than when the data are grouped by rhythm class. The results of the corresponding Tukey test place Italian and Japanese in one group, and Dutch, English and Spanish in another. This is because the %V of Spanish is closer to that of English and Dutch than to that of Italian, and likewise, the %V for Italian is closer to the %V of Japanese than to Spanish. These results are shown in Table 3 below.

[2] ΔV, ΔC and their standard deviations have been multiplied by 100 for ease of presentation.

Tukey HSD; α = 0.050 Q = 2.76833			
Language			Least Sq Mean
Japanese	A		50.036
Italian	A		47.657
Spanish		B	42.034
English		B	41.701
Dutch		B	39.606
Levels not connected by same letter are significantly different.			

Table 3. Mean %V by Language: Post-Hoc Tukey Test

%V was originally proposed as a useful measure for classifying – or at least differentiating – languages based on their perceived speech rhythm. A property we might expect from such a measure is consistency, not only across methodological replications of the original study, but also across speakers of a single language, within a single speaker's results, or across sentences in a given data set. The results from at least three separate studies (Ramus et al. 1999, White and Mattys 2007, and the present study), as summarized in Table 4, differ enough to raise questions about whether %V could serve as such a measure with the desired degree of predictability.

Language	Ramus et al. %V (SD)	White and Mattys %V (SD)	Renwick %V (SD)
English	40.1 (5.4)	38 (0.5)	41.701 (6.0)
Dutch	42.3 (4.2)	41 (1.2)	39.606 (4.9)
French	43.6 (4.5)	45 (0.5)	n/a
Spanish	43.8 (4.0)	48 (0.8)	42.034 (5.6)
Italian	45.2 (3.9)	n/a	47.657 (3.9)
Japanese	53.1 (3.4)	n/a	50.036 (4.3)

Table 4. Summary: Cross-experimental %V Results

Across all three studies, we see that the rank order of languages by %V remains the same (with the exception of English and Dutch in the present study). However, the precise values for a given language vary from study to study – for example, Ramus et al. found a %V for English of 40.1, while White and Mattys measured 38, and this study measured 39.606. This cross-study variation is not surprising; it might be expected, given differences in stimuli and possible methodological differences. However, if we turn to the differences between adjacently-ranked languages in each study (i.e. English vs. Dutch), we find that individual languages are spread apart to different degrees along the %V continuum in each study. For example, in the Ramus et al. (1999) results, the difference between Italian and Japanese is nearly 8%; in the present study, this difference is just over 2%, and in fact here, the two languages are not statistically separable. The large degrees of variation between adjacent languages across studies are illustrated below in Figure 1.

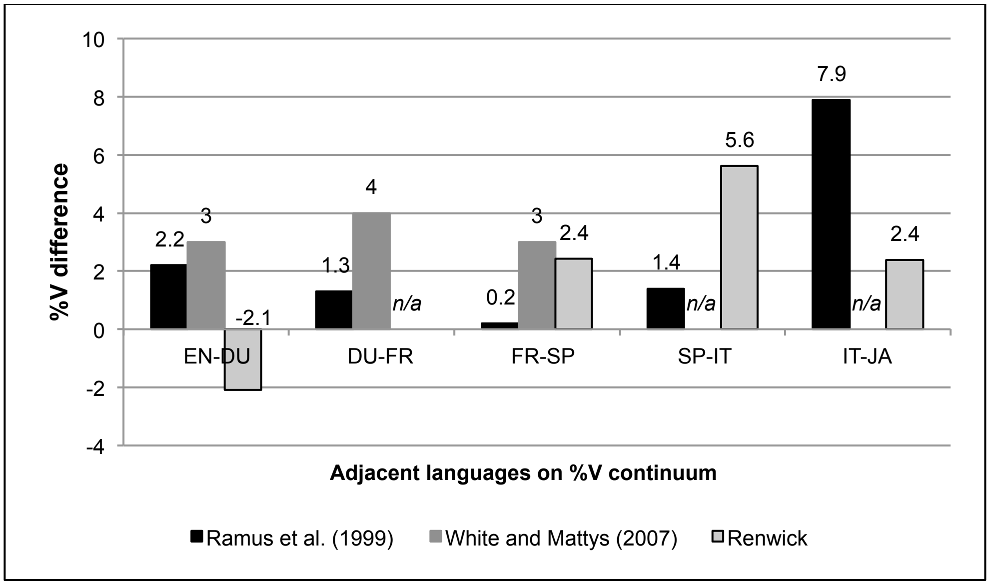

Figure 1. %V differences between languages adjacent on the %V scale, per study.

3 %V for individual utterances

Although the results in %V studies are typically presented as averages by language or by rhythm class, variability at the sentence level is arguably the type of information that listeners hear when judging or perceiving rhythm. To test how consistent this measure is across different utterances, Figure 2 plots %V on the vertical axis and the standard deviation of the duration of consonantal intervals (ΔC) horizontally; this technique has been used (Ramus et al. 1999) to show how rhythmically-similar languages cluster together based on these two durationally-based measurements. If this were true of the data collected here, we would expect the English and Dutch results to cluster where %V is low and ΔC is high, due to the large consonant clusters possible in these languages; Italian and Spanish, which allow some clusters of at most three consonants, would cluster at a medium %V and lower ΔC; and Japanese would have high %V but low ΔC, as a result of its restrictions on coda consonants. The results here, however, demonstrate something quite different: these measures are highly variable across the individual sentences whose %V values were earlier pooled for by-language analysis. While the sentences with lowest %V and highest ΔC are English, and those with the reverse properties are Japanese, we also find sentences in those languages that are adjacent in Figure 2.

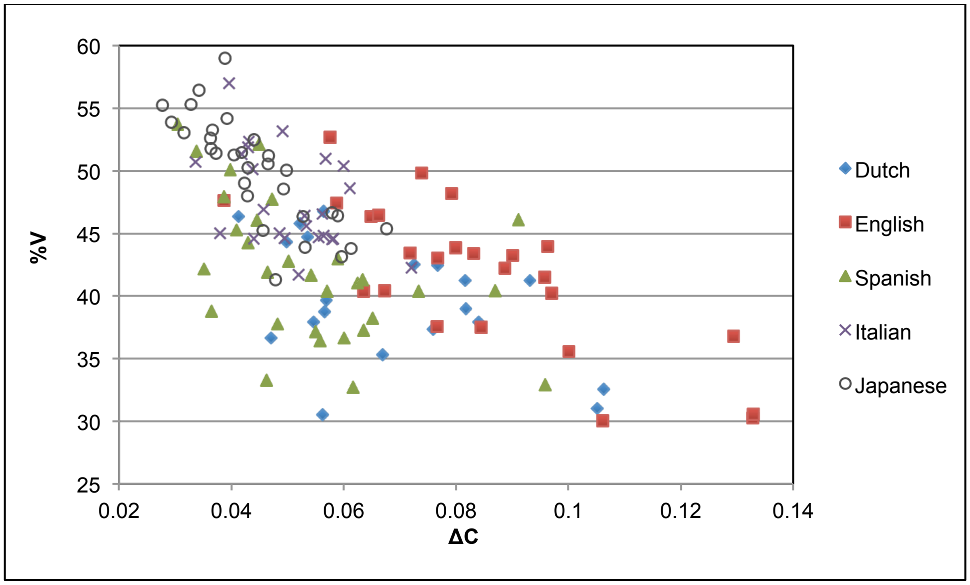

Figure 2. %V by ΔC, by Language: Individual Sentences

The high variability of %V in these data is again emphasized in Figure 3, which shows a breakdown of %V by language, using data pooled across speakers as in Table 1. While the mean %V values for the different languages correlate somewhat with previous %V results, two characteristics of the data are unexpected if %V is a reliable correlate of rhythm class. First, as we have seen in Table 1, both the mean %V of Spanish and its range are low, and comparable to that of English, which is in a different rhythmic category; and second, across languages, there is a great deal of overlap in %V values. For several English or Dutch sentences, for example, the %V is as high as the mean %V in Italian, and even the two most disparate languages, Dutch and Japanese, show substantial overlap (6 or more tokens out of 20).

These results add to the evidence that %V is a highly variable measure, across studies and even within individual speakers' data. The next section considers the question: Is there a phonological property with which %V correlates, that could help us refine our expectations for %V as an acoustic correlate of speech rhythm? To do this, we examine the results of Arvaniti (2009) and of White and Mattys (2007), as well as the phonotactic content of the data reported here.

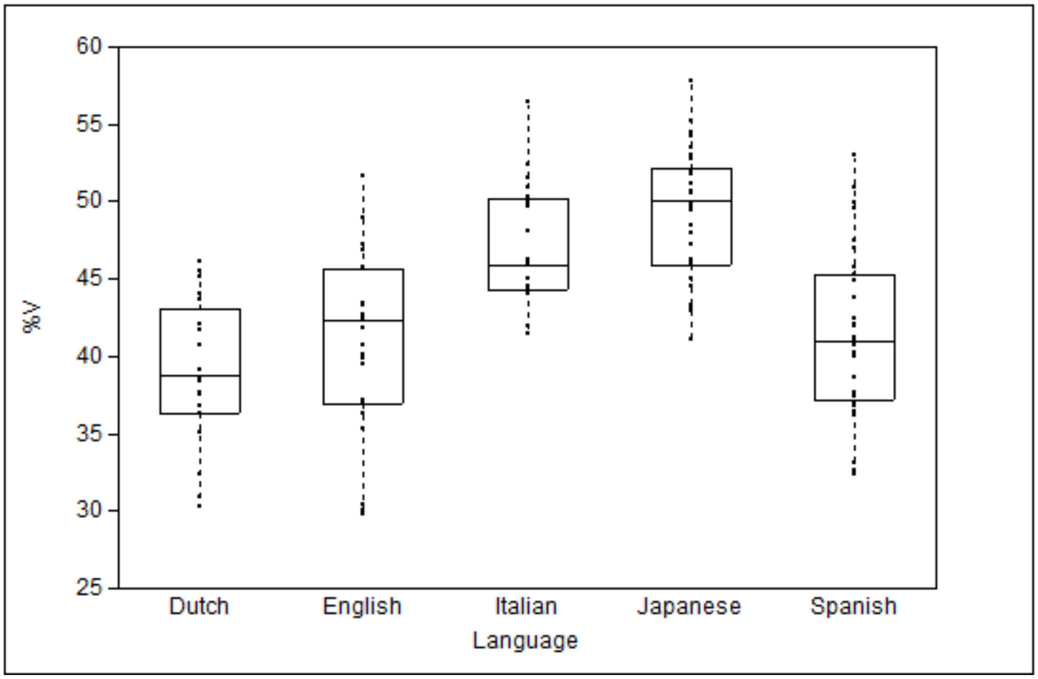

Figure 3. Variability and range of %V, by language: each dot represents the %V of an individual utterance.

4 The relationship between %V and syllable structure

The results presented here suggest that Spanish can pattern with English and Dutch in terms of %V, which is unexpected given the perceptual rhythm categorization results. Since other studies do not report statistical comparisons of %V across individual languages in addition to comparisons across the purported rhythm classes, we do not know how their data would group given the factor of language rather than rhythm class. However, the data presented here show considerable deviation from the results of Ramus et al. (1999), and especially from White and Mattys (2007), in which the %V of Spanish was 48% - higher even than the Italian result shown here. Given the stimulus conditions described in Section 2, it is plausible that the results for Spanish reported by White and Mattys were heavily influenced by the restrictions placed on segmental content in the stimuli, i.e. their stimuli included no approximants, and therefore very few coda consonants. It is also possible that the Spanish sentences analyzed in the present study happened to have an unusually large number of coda consonants, which may have lowered the overall %V result for the language. It appears, from Figure 3, that there is an enormous amount of variability in %V from one utterance to another in a language. Given the results of Arvaniti (2009) showing that variation in %V across groups of utterances is predicted by syllable structure, could the source of individual utterance variability lie in the segmental content, specifically the syllabic structure, of each individual utterance?

Arvaniti (2009) showed that ‘stress-timed’ stimuli designed to elicit reduced vowels and consonant clusters had the lowest %V and greatest ΔC, while ‘syllable-timed’ stimuli, which emphasized open syllables, had the inverse properties, and values for uncontrolled stimuli fell in between, regardless of language. This finding suggests that there is a link between syllable structure and %V, and that is what we investigate here. First, we show a breakdown of the data for each speaker, based on the proportions of segments in his or her utterances that belong to each major syllable position (onset, nucleus, coda); second, we analyze the statistical relationship between %V and syllable structure at the level of individual utterances.

The data used in this analysis are the sentences recorded for the %V replication in Section 2 (for list of stimuli, see Appendix A). The sentences, which contain between 15 and 21 syllables, were annotated for their lexical-form syllabic structure. This was done by considering the phonemic form of each sentence, and determining the syllabic structure of its words, ranging from a single vowel (V) to a syllable with both a complex onset and complex coda. In deciding what constituted a vowel or a consonant (and therefore a nucleus or marginal segment), the guidelines laid out for %V were followed as closely as possible. For example, syllabic sonorants like /l̩/ and /r̩/ (ɚ) of English were labeled as nuclei in words like 'neighbor' /nɛɪbɚ/. Glides were treated with a methodology similar to that of Ramus et al. (1999): pre-vocalic glides were grouped in the syllabic onset, but post-vocalic glides were grouped with the nucleus. Geminates, found word-medially in Italian and Japanese, were treated as both a coda and an onset, since they participate phonologically in both positions and are standardly analyzed as heterosyllabic.

We hypothesize that differences in %V are strongly influenced by the presence of coda segments, which not only raise the ratio of consonants to vowels in a particular stimulus sentence, but may also have the effect of shortening the duration of the preceding vowel. Thus we predict that the Italian and Japanese stimuli analyzed here, which have a higher %V value, will have many fewer coda segments than those in Spanish, Dutch and English. This prediction was tested by labeling the segments in the target utterances by syllable position – Onset, Nucleus and Coda – and calculating the contribution of each syllable position to the total duration of stimuli for each speaker. The results are shown as ratios in Figure 4, below. In this figure, the measure %Nuc is the same as %V, while its inverse '%C' is split into %Ons and %Coda. The duration of geminates, which are phonemic (i.e., contrastive with singleton consonants) in Italian and Japanese, were divided in half: 50% were ascribed to the coda, and 50% to the onset. In the case of geminate stops, only the closure was divided; 50% of the closure, plus the burst duration, went to the following onset.

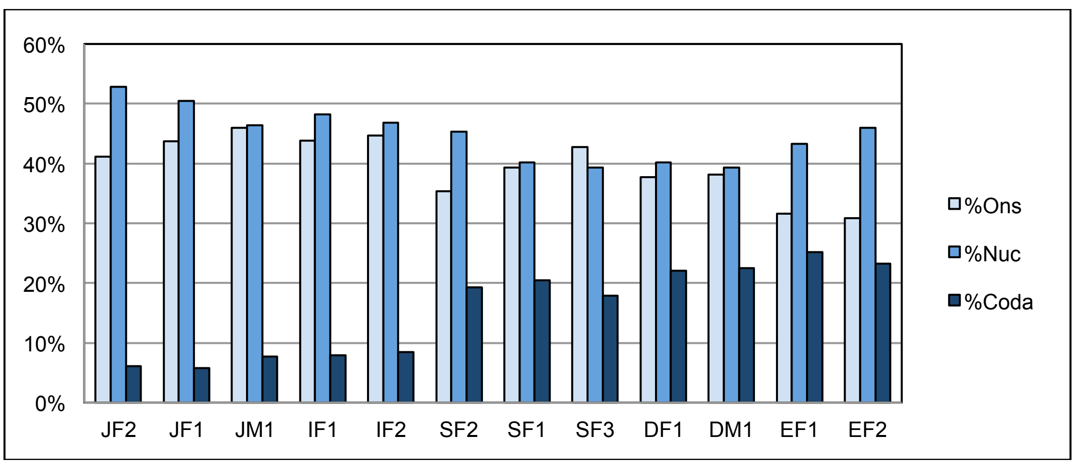

Figure 4. Ratio of Onset, Nucleus and Coda segments, by Speaker.

In Figure 4, we see that for speakers of Japanese and Italian, the durational proportion taken up by codas is very low; codas make up less than 10% of the total duration of the stimuli in those languages. By contrast, the data from Spanish, Dutch and English show higher coda proportions (18% - 25% of total duration), and a relatively lower proportion of nucleus duration (31% - 43%, vs. 41% - 46% for Italian and Japanese). Italian and Spanish are predicted to have similar coda measures on the basis of phonotactics, but it appears that the particular Spanish stimuli analyzed here happen to include more codas than expected, while there were very few in the Italian sentences (as in the Japanese). Thus the main result for the observations depicted in Figure 4 is that the presence of an unexpectedly high number of codas in the Spanish stimuli is the likely reason why our Spanish %V result differs greatly from those of other studies. Such a link between syllable structure and %V is not surprising; Ramus et al. (1999) propose a relationship of this type, which also reflects on the relevance of ΔC, in their original study:

> "Δ*C* and %*V* appear to be directly related to syllable structure. Indeed, a greater variety of syllable types means that some syllables are heavier […]. […] It is therefore not surprising to find English, Dutch and Polish (more than 15 syllable types) at one end of the Δ*C* and %*V* scales, and Japanese (4 syllable types) at the other. Thus, the nice fit between the (%*V*, Δ*C*) chart and the standard rhythm classes comes as an empirical validation of the hypothesis that rhythm contrasts are accounted for by differences in the variety of syllable structures" (1999: 273-274).

The findings of Figure 4, however, reinforce the volatility and lack of reliability inherent in the %V measure when it is applied to small corpora, which are subject to random variation in syllable structure. The stimuli used in the original %V study (Ramus et al. 1999) as well as in this replication were controlled only for the number of syllables in each sentence; they were not controlled for syllable types or segmental content. Across the existing %V studies, no set of stimuli is claimed to be representative of the frequencies with which different segments and syllable structure types appear in each language tested, so it is difficult to test the possibility of a consistent relationship among the phonotactics of a language, its perceptual rhythmic classification, and a reported %V value based on a particular speech sample which is uncontrolled for how well it represents the syllable structure of the language.

We have shown that the discrepancies in Spanish %V results across studies may well be a function of the syllabic structures in the recordings analyzed in each study. We have used the differing ratios of Onset, Nucleus and Coda to make this point. We now examine the relationship between segmental structure and %V at the level of individual utterances in greater detail, by modeling %V as a function of syllable structure (open vs. closed, i.e. with and without coda, and with vs. without onsets).

Once each sentence was syllabified, its open and closed syllables were counted. An open syllable is one without a coda, minimally (V); a closed syllable is one with a coda, minimally (VC). The number of open syllables was divided by the total number of syllables in the sentence, to create a ratio: %Open, or the percentage of open syllables in the sentence. Additionally, the numbers of complex onsets and complex codas in each sentence were counted, for inclusion in modeling. The motivation behind this was that %Open only took into account the presence (or lack thereof) of at least one coda segment; however, we expect %V to decrease further as complex clusters increase in number, so these measures could prove to be significant in predicting %V.

To quantify the relationship between syllable structure and %V, a statistical analysis was conducted. A simple ANOVA showed that the relationship between %V and %Open is quite strong ($r = .710$), as seen by the line of fit in Figure 5, and %Open was a highly significant predictor of %V ($p \sim 0.0$).

Figure 5 shows the relationship between %V, on the Y-axis, and %Open (the number of open syllables in a particular utterance divided by the total number of syllables), on the X-axis. Each data point represents a single utterance; all 129 utterances analyzed for the results in Table 1 are included. Figure 5 shows a clear trend: The higher the %Open value of a sentence, the higher its %V, and vice-versa. This correlates with the fact that %Open is a significant factor in accounting for %V, shown below.

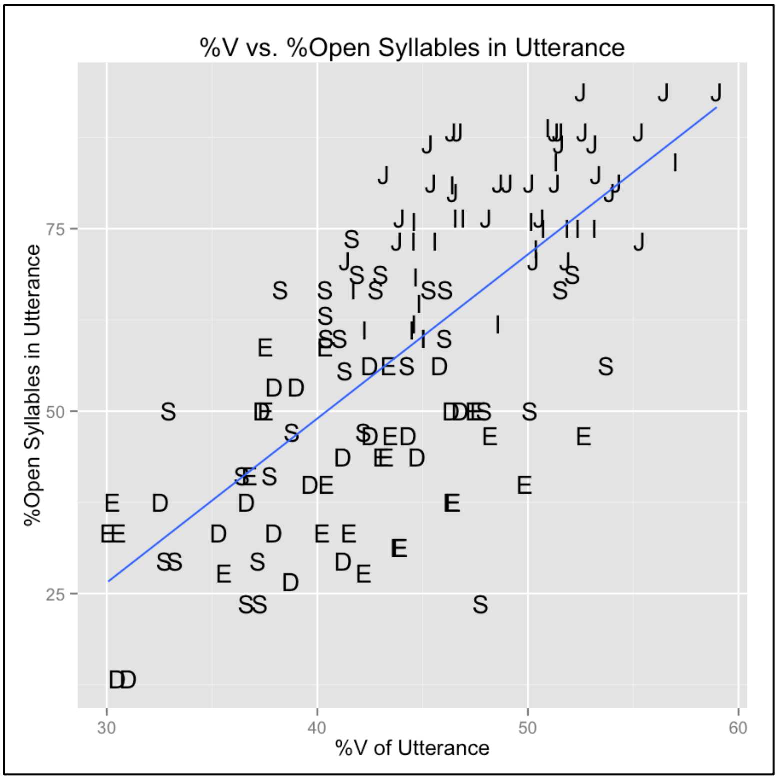

Figure 5. %Open vs. %V. Each letter represents one utterance. **D** = Dutch; **E** = English; **I** = Italian; **J** = Japanese; **S** = Spanish. Blue line indicates correlation between %Open and %V (r = 0.710).

In order to test the relationship between the quantifiable aspects of syllable structure and the variability in %V, a mixed-effects model was constructed. %V remained the dependent variable; the fixed effects were %Open and the number of complex onsets in each sentence; and Speaker was included as a random effect. In this model, %Open was a significant predictor of %V (F(1, 73.02) = 60.7593, $p < 0.0001$), as was the presence of a complex onset (F(1, 124.3) = 5.1534, p = 0.0249). The counterpart to Complex Onsets – Complex Codas – was found to be an insignificant predictor of %V in a model that also included %Open. This was presumably due to the collinearity between the two variables. However, when tested separately, the number of complex codas correlates reasonably well, inversely, with the %V value of a sentence (r = -0.44). The number of complex onsets is also negatively related to %V (r = -0.37). Neither of these predictors correlates to %V as well as %Open.

Two other statistical models, based on the mixed effects model above, were tested, but were found to have insignificant predictors. These models were nearly identical to the mixed effects model proposed (including Speaker as a random factor), but each had a single specific difference: one also included Rhythm Class as a fixed effect, and the other included Language as a fixed effect. In both cases, this extra factor was found to be insignificant (F(2, 12.61) = 0.1040, p = 0.9020 for Rhythm Class, F(4, 8.263) = 0.2270, p = 0.9159 for Language). In models that included an interaction between Rhythm Class or Language and %Open, this interaction also failed to reach significance. This result is different from the simple ANOVA models of %V in Section 2, in which post-hoc Tukey tests found significant differences across rhythm classes or languages, presumably due to the fact that the simple ANOVAs did not take syllable structure into account. That is, when we do use syllable structure to model %V, it turns out to be a better predictor than either Language or Rhythm Class.

In sum, these results indicate that when syllable structure is taken into account for predicting %V, additional knowledge about language or rhythm classification does *not* contribute additional predictive power. Thus the phonotactic properties of a sentence are sufficient to predict %V, which these data do not show to be otherwise related to the rhythm class or language in which that sentence is produced.

5 Discussion and Conclusions

The previous section used two measures to quantify the relationship between %V and syllable structure. First, we demonstrated that in this set of utterances, languages with lower %V values (Spanish, English, Dutch) tended to have a greater durational proportion of coda consonants than those with higher %V (Italian, Japanese). Second, we found significant relationships between %V and phonotactics, in the form of a statistical model taking into account the percentage of open syllables in a given utterance, as well as the number of complex onsets it contained. This model accounted for %V so well that additional factors for the language in which an utterance was spoken, or its purported rhythm class, were not significant predictors. Taken together, these results cast doubt on the usefulness of %V as an independent predictor of timing-related rhythm distinctions, while indicating that the insights about syllable structure that partially inspired the development of this measure are by themselves quite powerful determinants of this value. The experiments described above illustrate the difficulty in replicating the original %V results with the degree of precision we might hope for in order to pinpoint divisions between rhythm classes. Additionally, the results presented here have shown why %V is so variable, through analyses that require more detailed analysis of the syllable structure of the sampled utterances.

It is possible that the small sample sizes used in this experiment influenced the lack of significance of factors like Rhythm and Language; given a large-enough sample, it is possible that the data from a given language or typological phonotactic class would cluster in a particular region of the %V – %Open space. Hints of this are visible already in Figure 5; most of the sentences with high %V and %Open are Japanese and Italian, while most of those with low %V and %Open are Dutch and English. However, the overlap of the Spanish utterances with those in Dutch and English reminds us that %V is strongly dependent on the phonotactics of each sentence being measured, and that it may not capture any information independently of those facts. We must also recall that the sentences used in these studies should not be considered representative of the phonotactic structure of a given language, particularly where the relative frequencies of different syllable structures are concerned. The stimuli used may over-sample some forms, for example through repetition of certain words, and under-sample others. Furthermore, any sample-size problems are not limited to the present study: the number of sentences analyzed per language here was at least equal to the number included in the original %V study (Ramus et al. 1999), and for all languages except Dutch, this study in fact analyzed more utterances than the original study.

In sum, the results presented in this paper emphasize two aspects of the %V measure. First, it is volatile across utterances in a language, because of its dependence on the phonotactic content of each analyzed sentence. Second, at least when estimated from small corpora, %V does not provide additional information about the rhythm class of a language that cannot be gleaned from phonotactic facts about the utterances involved, such as the presence of syllable codas and complex onsets. In the data examined here, when these factors are taken into account, neither the language of a sample nor its rhythmic category contributes significantly to a model of %V variability. It is possible that analysis of a larger sample, which accurately reflects the distributional and token frequency facts about the syllable structure characteristics of a language, would allow statistical modeling in which information about rhythm class could be predictive, or would capture the effects of certain phonological characteristics (i.e. vowel reduction) on %V. Additionally, we should keep in mind that while the utterances analyzed in %V studies are typically recorded in a careful-speech setting, more casual speech might have different characteristics or introduce additional sources of variability; and yet another source of variability to consider are intrinsic durational differences between segments, which directly factor into an utterance's %V. While the data analyzed here suggest that the nature of the sample determines almost entirely the %V result, we cannot exclude the idea that %V could help us reach more useful conclusions, if we had more information about the phonotactics of each language and their relative distributions, and a sufficient sample size.

Proponents of rhythm-class distinctions might argue that these timing-related differences should be measurable regardless of the segmental content of a particular utterance, and so it is possible that %V is inadequate because a division more fine-grained than a simple consonant-vowel distinction is necessary. Other investigators have chosen different dividing lines for calculating timing ratios, such as Dellwo et al. (2007), who show that measuring voiced vs. voiceless intervals in natural speech provides a clear separation between rhythm classes, when graphed on the plane of %Voiced vs. Variability of Voiceless Intervals (analogous to %V vs. ΔC). Dellwo et al. (2007) argue that the distinction of "vocalic" and "consonantal" intervals is the wrong distinction, i.e. that it may be too complex for infants to perceive, and that the crucial perceptual information may be even more basic. However, Nespor et al. (2003) have argued that the V/C distinction is perceptually highly reliable, and is backed up by demonstrable psychological reality. Other attempts to develop an acoustic-phonetic measure that reflects rhythmic differences, such as that proposed by Steiner (2003), have taken sonority into account. Steiner annotated utterances based on their segments' phonological sonority level and argued that some consonant classes bear more of the functional load for determining rhythm class: "[T]he statistical distribution of certain significant consonant classes in individual languages determines these languages' membership in rhythmic classes" (Steiner 2003: 6). She concluded that phonotactics and syllable structure play a large role in determining perceived rhythm. The results shown here support that claim, which echoes Dauer (1983) and Auer (1990). Related pilot studies performed by the present author on the relationship between sonority (via intensity, found by Parker 2002 to be a good correlate of sonority) and rhythm class included repeated attempts to find a cut-off point between sonorous and non-sonorous material so that their relative proportions would reflect a rhythm-class divide. However, these studies obtained only null results, and an acoustic correlate linking rhythm and sonority remains elusive.

Taken together, the lines of work summarized here continue to demonstrate that, despite the fact that both users and scholars of language have powerful intuitions about differences in rhythm across languages, it remains an open question whether those differences can be laid entirely at the feet of phonotactic structure, or whether there are some additional phonetic implementation differences across languages that might underlie these perceived rhythmic differences.

References

Abercrombie, D. (1965). *Studies in phonetics and linguistics*. London: Oxford University Press.

Arvaniti, A. (2009). Rhythm, Timing and the Timing of Rhythm. *Phonetica, 66*, 46-63.

Auer, P. (1990). "Stress-Timing" vs. "Syllable-Timing" from a Typological Point of View. In B. Palek (Ed.), *Proceedings of the Conference Linguistics and Phonetics: prospects and applications* (pp. 292-305). Prague: Charles University Press.

Dauer, R. M. (1987). Phonetic and phonological components of language rhythm. In *Papers presented at the 11th International Congress of Phonetic Sciences* (Tallinn), 447-450.

Dauer, R. M. (1983). Stress-timing and syllable-timing reanalyzed. *Journal of Phonetics, 11*, 51-62.

Dellwo, V. (2006). Rhythm and speech rate: A variation coefficient for deltaC. In P. Karnowski, & I. Szigeti (Eds.), *Language and language processing: Proceedings of the 38th Linguistic Colloquium* (pp. 231–241). Frankfurt: Peter Lang.

Dellwo, V., Fourcin, A., & Abberton, E. (2007). Rhythmical classification of languages based on voice parameters. In *Papers Presented at the 16th International Congress of Phonetic Sciences* (Saarbrücken), 1129-1132.

Dellwo, V., & Wagner, P. (2003). Relations between language rhythm and speech rate. In *Papers Presented at the 15th International Congress of Phonetic Sciences* (Barcelona), 471-474.

James, L. A. (1940). *Speech signals in telephony*. London: Pitman & Sons.
Lehiste, I. (1973). Rhythmic units and syntactic units in production and perception. *Journal of the Acoustical Society of America, 54*, 1228-1234.
Lehiste, Ilse. (1977). Isochrony reconsidered. *Journal of Phonetics, 5*, 253-263.
Nazzi, T., Bertoncini, J., & Mehler, J. (1998). Language Discrimination by Newborns: towards an Understanding of the Role of Rhythm. *Journal of Experimental Psychology: Human Perception and Performance, 24*, 756-766.
Nespor, M. (1990). On the rhythm parameter in phonology. In I. M. Roca (Ed.), *Logical issues in language acquisition* (pp. 157-175). Dordrecht: Foris.
Nespor, M., Peña, M., & Mehler, J. (2003). On the different roles of vowels and consonants in speech processing and language acquisition. *Lingue e linguaggio, 2*, 203-229.
Parker, S. G. (2002). *Quantifying the Sonority Hierarchy*. Amherst, MA: University of Massachusetts Department of Linguistics.
Port, R. F., Dalby, J. & O'Dell, M. (1987). Evidence for mora-timing in Japanese. *Journal of the Acoustical Society of America, 81*, 1574-1585.
Ramus, F., & Mehler, J. (1999). Language identification with suprasegmental cues: A study based on speech resynthesis. *Journal of the Acoustical Society of America, 105*, 512-521.
Ramus, F., Nespor, M., & Mehler, J. (1999). Correlates of Linguistic Rhythm in the Speech Signal. *Cognition, 73*, 265-92.
Ramus, F., Dupoux, E., & Mehler, J. (2003). The psychological reality of rhythm classes: perceptual studies. *Proceedings of the 15th International Congress of Phonetic Sciences,* (Barcelona), 337-342.
Roach, P. (1982). On the distinction between ‘stress-timed’ and ‘syllable-timed’ languages. In D. Crystal (Ed.), *Linguistic controversies*. London: Edward Arnold.
Steiner, I. (2003). A refined acoustic analysis of speech rhythm. *Proceedings of the 38th Linguistic Colloquium* (Budapest).
Toro, J. M., Trobalon, J. B., & Sebastián-Gallés, N. (2003). The use of prosodic cues in language discrimination tasks by rats. *Animal Cognition, 6*, 131-136.
White, L., & Mattys, S. L. (2007). Calibrating rhythm: First language and second language studies. *Journal of Phonetics, 35,* 501-522.

Department of Linguistics
Cornell University
Ithaca, NY 14853
mer56@cornell.edu

APPENDIX A: Experimental Stimuli

Reproduced from Nazzi et al. (1998)

Dutch Stimuli

1. De jongen stond vroeg op om in zijn nieuwe boek te lezen.
2. De boze demonstranten raakten slaags met de politie.
3. Het omstreden artikel zorgde voor heel wat opschudding.
4. De prinses had kramp in haar hand van het lintjes doorknippen.
5. Erwtensoep met worst is nog steeds zijn favoriete gerecht.
6. De uitgever spande een proces aan tegen de schrijver.
7. De dader werd helaas bij gebrek aan bewijs vrijgesproken.
8. Het belang van milieubewustheid wordt steeds vaker ingezien.
9. Zij heeft voor alles altijd een psychologische verklaring.
10. De geur van vers gezette koffie verspreidde zich door het huis.
11. Onze laatste aanwinst is een prachtige antieke sofa.
12. Als je nog eens hier komt zwemmen, neem dan vooral een handdoek mee.
13. Op het ijs stond een kraampje met chocolademelk en stroopwafels.
14. Nauwelijks was het concert uit of een daverend applaus brak los.
15. Delegaties uit meer dan twintig landen komen naar dit congres.
16. Het lawaai van de machines maakte elk gesprek onmogelijk.
17. Een gevoel van enornie opluchting maakte zich van hem meester.
18. Het draaiorgel is bijna helemaal uit het straatbeeld verdwenen.
19. In tegenstelling tot zijn broer heeft hij altijd van schaken gehouden.
20. Beeldend kunstenaars doen vaak hun inspiratie op in grote steden.

21. Het economisch klimaat is niet gunstig voor het vinden van een baan.
22. Onder grote spanning vertonen de meeste mensen hun ware aard.
23. In het oude centrum van de stad vind je nog middeleeuwse huizen.
24. Door het uitvallen van de microfoons was de toespraak onverstaanbaar.
25. In die dierentuin is voor de eerste maal een pandabeertje geboren.
26. De favoriete werkplek van de schrijfster was een oude villa aan zee.
27. Niemand heeft ooit kunnen achterhalen waar het geld terechtgekomen is.
28. Het was die dag zo heet dat de toeristen spontaan in de fontein sprongen.
29. De jonge architect leverde een verbluffend staaltje van vakmanschap.
30. Het restaurant werd terstond gesloten om redenen van hygiene.
31. Volgens oud gebruik wordt in dat dorp ieder jaar een uitbundig oogstfeest gevierd.
32. Honderdveertigduizend soldaten worden in allerijl gemobiliseerd.
33. Plotseling realiseerde ze zich dat ze het zwervende leven moe was.
34. De reizigers ontdekten tot hun schrik dat de trein op een ander perron stond.
35. Na maanden van discussie is nu eindelijk een standbeeld geplaatst in het park.
36. Het nieuwe model fiets werd afgedaan als een vergezocht modeverschijnsel.
37. Het kind bracht zijn ouders tot schateren met zijn imitatie van de lerares.
38. De nieuwe concertzaal lijkt meer op een ouderwetse fabriekshal dan op iets anders.
39. Dankzij de volle inzet van alle medewerkers is het project een succes.
40. Ondanks het radiobericht waren de mensen niet voorbereid op de orkaan.

English Stimuli

1. My grandparents' neighbor is the most charming person I know.
2. The local train left the station more than five minutes ago.
3. The next local elections will take place during the winter.
4. Much more money will be needed to make this project succeed.
5. The art gallery in this street was opened only last week.
6. A hurricane was announced this afternoon on the TV.
7. The parents quietly crossed the dark room and approached the boy's bed.
8. The first flowers have bloomed due to the exceptional warmth of March.
9. In this famous coffee shop you will eat the best doughnuts in town.
10. The young boy got up quite early in order to watch the sunrise.
11. This supermarket had to close due to economic problems.
12. The committee will meet this afternoon for a special debate.
13. Nobody noticed when the children slipped away just after dinner.
14. In this case, the easiest solution seems to appeal to the high court.
15. The last concert given at the opera was a tremendous success.
16. Science has acquired an important place in western society.
17. This rugby season promises to be a very exciting one.
18. Artists have always been attracted by the life in the capital.
19. Finding a job is difficult in the present economic climate.
20. Trade unions have lost a lot of their influence during the past ten years.
21. The library is opened every day from eight A.M. to six P.M.
22. They didn't hear the good news until last week on their visit to their friends.
23. Most European banks close extremely early on Friday afternoons.
24. Having a big car is not something I would recommend in this city.
25. This year's Chinese delegation was not nearly as impressive as last year's.
26. The city council has decided to renovate the medieval center.
27. There is an important market twice a week on the main square of the village.
28. The government is planning a reform of the educational program.
29. No welcome speech will be delivered without the press officer's agreement.
30. The recent rainfall has caused very severe damage in the higher valleys.
31. The woman over there is an eminent specialist in plastic surgery.
32. Seven paintings of great value have recently been stolen from the museum.
33. The Green Party has unexpectedly gained strong support from middle class people.
34. This is the first time an international exhibition takes place in this town.

35. Mothers usually leave the maternity unit two days after giving birth.
36. The rebuilding of the city started the very first day after the earthquake.
37. Most of the supporters of the football club had to travel for an entire day.
38. In spite of technical progress, predicting the weather is still very difficult.
39. The latest events have caused an outcry in the international community.
40. It is getting very easy nowadays to find a place in a nursery school.

Italian Stimuli

1. Il mercato si tiene sulla piazza ogni giorno.
2. La moglie del farmacista sa sempre ciò che vuole.
3. Le strade che danno sulla piazza sono bloccate.
4. Le forti piogge della primavera sono dannose.
5. II bambino scese prestissimo per vedere l'alba.
6. Il teatro ha introdotto molte nuove discipline.
7. La stagione musicale non offrirà grandi novità.
8. Un quadro molto famoso è stato mal restaurato.
9. Non ha mai voluto rendersi conto dei suoi gran difetti.
10. La radio ha comunicato questa mattina la notizia.
11. La ricostruzione della città dovrà farsi lentamente.
12. L'ozio non è il solo padre dei gran vizi dell'umanità.
13. Tutte le deroghe devono recare prova di conformità.
14. Il vantaggio di poter scrivere liberamente è immenso.
15. Il venerdì sera le banche chiudono sempre con anticipo.
16. La situazione della bilancia dei pagamenti non mi lascia mai tranquillo.
17. I sostenitori della riforma si sono trovati sulla piazza principale.
18. Credo che riuscirai nei tuoi piani senza farti problemi di sorta.
19. La statistica permette di comprendere la scienza sperimentale.
20. Il giunto meccanico è troppo debole per sopportare quel peso.

Japanese Stimuli

1. Bankokuhakurankai wa sakunen kasaisareeta.
2. Monku wa shihainin ni iuno ga tettoribayai.
3. Oono shigo ni machi no saiken ga hajimatta.
4. Noomin no sonchoo ni taisuru fuman ga tamatta.
5. Haru no koozui de zuibun ookina higaiga deta.
6. Shussango sooki ni taiinsuru keekooga tsuyomatta.
7. Konopanya no keiki wa konokaiwai de hyoobanda.
8. Konshuu mo terebibangumi o mirujikan ga nai.
9. Kochira no kata wa keiseigeka no senmonka desu.
10. Tsugino chihoosenkyo wa kondo no harugoro deshoo.
11. Saikin no jiken de sekai no yoron wa konranshiteiru.
12. Kaikakusuishinha ga kenchoomae de demokooshinshita.
13. Kesa no rajiode taifuukeihoo ga hatsureisareta.
14. Guusuuno hini wa kono hiroba ni miseya ga deru.
15. Tsugi no gekijooshiizun wa totemo kyoomibukaidaroo.
16. Kusuriya no kamisan wa moosugu kaimononi deru.
17. Tokurei wa kaino sanseinashini wa mitomerarenai.
18. Moosugu rinjikokkaino kaiki ga hajimaruhazuda.
19. Kookyooko otsuukikan no seibiwa doomitemo fujuubunda.
20. Boku wa baiorin no keiko o kazoekirenaikurai yasunda.
21. Hinode o mirutame ni sonoko wa hayaku kishooshita.
22. Kokono shokudoo wa eiseijoo no mondai de heisasareta.
23. Moo gofun ijoo mae ni tokkyuu wa hoomu ni tsukimashitayo.
24. Jikai no kaikaku no taishoo wa gakkookyooiku no naiyoodesu.
25. Kokosuunen de roodoodantai no eikyooryoku ga teikashita.
26. Koocha demo nominagara koko de matasete morauyo.

27. Keikanwa yoogishano fuuteini nita shooni no mikaketa.
28. Sobo no sumai no gokinjo wa yoi hitotachi bakaridesu.
29. Keekaku no jitsugen niwa shikin ga kanari hitsuyoodeshoo.
30. Chuusankaikyuunaibu de kankyohogoha ga seiryoku o nobashita.
31. Kotoshi no nihondaihyoodan wa kyonenyori ninzuu ga sukunai.
32. Doo gijutsu ga shinposhitemo tenkiyohoo wa tekichuushinai.
33. Daitoshi no seikatsu wa tasuu no geijutsuka o hikitsuketa.
34. Dokoka tooku de yakiimoya no fuenone ga natteiruyooda.
35. Oozeino kyakunomae de tabako o suuno wa yametahoo ga yoi.
36. Ryooshin wa mono oto o tatezu ni kodomo no soba ni chikazuita.
37. Kodomo o kooritsukoo ni susumaseru nowa muzukashikunai.
38. Konokeeki no jootainomama dewa shokusagashi wa taihenda.
39. Shitookyoku ga rekishikuiki no saikaihatsu ni chakushushita.
40. Amerikajin ga gaikokujin dato kangae tara kookaisuruzo.

Spanish Stimuli
1. La radio annunció esta noticia el miércoles.
2. El niño se levantó temprano para ver el sol.
3. El ladron se fue con el reloj de oro de mi padre.
4. Esta panadería hace los mejores pasteles.
5. Los padres se acercaron del niño sin hacer ruido.
6. Los bancos cierran particularmente temprano el viernes.
7. Los artistas siempre fueron atraídos por las ciudades.
8. Los recientos acontecimientos hicieron escandalo.
9. La reconstrucción de la ciudad empezó el año pasado.
10. Encontrar un empleo no es facil en el contexto actual.
11. Las madres salen cada vez mas rápido de la maternidad.
12. La última exposición universal fue en San Francisco.
13. La pintura moderna tiene un éxito cada vez más grande.
14. El presupuesto del ministerio de la cultura bajo mucho.
15. La Consagración de la Primavera hizo escándalo en París.
16. Esos leones lo pasaron regio comiéndose a unos cristianos.
17. La corriente ecológica creció bastante en la clase media.
18. En verano, las grandes ciudades europeas se llenan de turistas.
19. Si a los tontos les crecieran alas, no se vería nunca más el sol.
20. Dios, fijándose que no podía vigilar a todos, creó las madres.

APPENDIX B: %V Replication Data

			V-Intervals (s.ms)		C-Intervals (s.ms)		Total for Sentence		
Language	Speaker	Sentence[3]	Duration	Std. Dev.	Duration	Std. Dev.	Duration (s.ms)	%V	Rhythm Class
Dutch	DF1	1	1.107	0.043	1.684	0.057	2.791	39.663	Stress
Dutch	DF1	2	1.14	0.043	1.783	0.082	2.923	39.001	Stress
Dutch	DF1	4	0.816	0.027	1.86	0.056	2.676	30.493	Stress
Dutch	DF1	5	1.011	0.030	1.851	0.067	2.862	35.325	Stress
Dutch	DF1	6	1.148	0.048	1.443	0.050	2.591	44.307	Stress
Dutch	DF1	7	1.347	0.049	1.594	0.052	2.941	45.801	Stress
Dutch	DF1	8	1.122	0.039	1.941	0.047	3.063	36.631	Stress
Dutch	DF1	9	1.169	0.048	1.444	0.054	2.613	44.738	Stress
Dutch	DF1	10	1.129	0.053	1.893	0.076	3.022	37.359	Stress
Dutch	DF1	11	1.396	0.061	1.615	0.041	3.011	46.363	Stress

[3] Sentence number corresponds to the order in Appendix A, above.

			V-Intervals (s.ms)		C-Intervals (s.ms)		Total for Sentence		
Language	Speaker	Sentence[3]	Duration	Std. Dev.	Duration	Std. Dev.	Duration (s.ms)	%V	Rhythm Class
Dutch	DM1	2	1.015	0.039	1.66	0.084	2.675	37.944	Stress
Dutch	DM1	3	1.008	0.034	1.594	0.057	2.602	38.739	Stress
Dutch	DM1	4	0.935	0.026	2.078	0.105	3.013	31.032	Stress
Dutch	DM1	5	1.078	0.035	1.764	0.055	2.842	37.931	Stress
Dutch	DM1	6	1.16	0.063	1.569	0.072	2.729	42.506	Stress
Dutch	DM1	7	1.255	0.042	1.699	0.077	2.954	42.485	Stress
Dutch	DM1	8	1.024	0.030	2.122	0.106	3.146	32.549	Stress
Dutch	DM1	9	1.176	0.040	1.677	0.093	2.853	41.220	Stress
Dutch	DM1	11	1.348	0.052	1.533	0.056	2.881	46.789	Stress
Dutch	DM1	14	1.351	0.049	1.925	0.082	3.276	41.239	Stress
English	EF1	1	1.187	0.066	1.764	0.097	2.951	40.224	Stress
English	EF1	2	1.324	0.081	1.722	0.072	3.046	43.467	Stress
English	EF1	3	0.907	0.038	2.11	0.106	3.017	30.063	Stress
English	EF1	4	1.245	0.050	1.835	0.067	3.080	40.422	Stress
English	EF1	7	1.567	0.060	2.075	0.077	3.642	43.026	Stress
English	EF1	8	1.108	0.038	2.554	0.133	3.662	30.257	Stress
English	EF1	9	1.78	0.064	2.059	0.065	3.839	46.366	Stress
English	EF1	10	1.447	0.048	1.853	0.080	3.300	43.848	Stress
English	EF1	12	1.054	0.048	1.753	0.077	2.807	37.549	Stress
English	EF1	13	1.199	0.046	1.996	0.084	3.195	37.527	Stress
English	EF1	19	1.488	0.036	1.635	0.039	3.123	47.646	Stress
English	EF1	20	1.325	0.040	2.399	0.100	3.724	35.580	Stress
English	EF2	1	1.308	0.069	1.844	0.096	3.152	41.497	Stress
English	EF2	2	1.675	0.080	1.8	0.079	3.475	48.201	Stress
English	EF2	3	1.083	0.028	2.462	0.133	3.545	30.550	Stress
English	EF2	5	1.517	0.061	1.527	0.074	3.044	49.836	Stress
English	EF2	6	1.619	0.071	1.455	0.058	3.074	52.668	Stress
English	EF2	7	1.786	0.075	2.343	0.090	4.129	43.255	Stress
English	EF2	9	1.869	0.075	2.155	0.066	4.024	46.446	Stress
English	EF2	10	1.668	0.052	2.129	0.096	3.797	43.929	Stress
English	EF2	11	1.425	0.060	1.859	0.083	3.284	43.392	Stress
English	EF2	16	1.476	0.065	2.536	0.129	4.012	36.790	Stress
English	EF2	17	1.334	0.037	1.97	0.063	3.304	40.375	Stress
English	EF2	19	1.724	0.052	1.912	0.059	3.636	47.415	Stress
English	EF2	20	1.746	0.053	2.391	0.089	4.137	42.204	Stress
Spanish	SF1	1	0.954	0.031	1.409	0.057	2.363	40.372	Syllable
Spanish	SF1	2	0.915	0.022	1.314	0.062	2.229	41.050	Syllable
Spanish	SF1	3	1.063	0.035	1.339	0.043	2.402	44.255	Syllable
Spanish	SF1	4	0.996	0.055	1.382	0.046	2.378	41.884	Syllable
Spanish	SF1	5	1.048	0.027	1.139	0.039	2.187	47.920	Syllable
Spanish	SF1	6	0.915	0.018	1.837	0.046	2.752	33.249	Syllable
Spanish	SF1	7	0.973	0.031	1.698	0.056	2.671	36.428	Syllable
Spanish	SF1	8	0.961	0.015	1.517	0.036	2.478	38.781	Syllable
Spanish	SF1	9	1.15	0.026	1.389	0.041	2.539	45.293	Syllable
Spanish	SF1	10	0.93	0.019	1.609	0.060	2.539	36.629	Syllable
Spanish	SF2	1	0.877	0.024	1.173	0.050	2.050	42.780	Syllable

			V-Intervals (s.ms)		C-Intervals (s.ms)		Total for Sentence		
Language	Speaker	Sentence[3]	Duration	Std. Dev.	Duration	Std. Dev.	Duration (s.ms)	%V	Rhythm Class
Spanish	SF2	2	0.939	0.019	1.101	0.045	2.040	46.029	Syllable
Spanish	SF2	3	1.173	0.036	1.011	0.030	2.184	53.709	Syllable
Spanish	SF2	4	1.161	0.101	1.067	0.045	2.228	52.110	Syllable
Spanish	SF2	5	1.289	0.033	1.284	0.040	2.573	50.097	Syllable
Spanish	SF2	6	1.003	0.018	1.695	0.055	2.698	37.176	Syllable
Spanish	SF2	7	1.004	0.037	1.656	0.048	2.660	37.744	Syllable
Spanish	SF2	8	0.942	0.021	1.291	0.035	2.233	42.185	Syllable
Spanish	SF2	9	1.153	0.033	1.084	0.034	2.237	51.542	Syllable
Spanish	SF2	10	1.479	0.056	1.618	0.047	3.097	47.756	Syllable
Spanish	SF3	2	1.217	0.037	1.794	0.087	3.011	40.418	Syllable
Spanish	SF3	4	1.021	0.059	1.353	0.059	2.374	43.008	Syllable
Spanish	SF3	6	1.13	0.052	2.322	0.062	3.452	32.735	Syllable
Spanish	SF3	9	1.403	0.067	1.642	0.091	3.045	46.076	Syllable
Spanish	SF3	10	1.178	0.053	1.983	0.063	3.161	37.267	Syllable
Spanish	SF3	11	1.3	0.022	1.848	0.063	3.148	41.296	Syllable
Spanish	SF3	12	1.044	0.044	2.126	0.096	3.170	32.934	Syllable
Spanish	SF3	13	1.257	0.020	1.855	0.073	3.112	40.392	Syllable
Spanish	SF3	14	1.355	0.050	1.898	0.054	3.253	41.654	Syllable
Spanish	SF3	16	1.33	0.065	2.147	0.065	3.477	38.251	Syllable
Italian	IF1	1	1.155	0.040	1.41	0.038	2.565	45.029	Syllable
Italian	IF1	3	1.334	0.042	1.592	0.053	2.926	45.591	Syllable
Italian	IF1	4	1.476	0.050	1.343	0.043	2.819	52.359	Syllable
Italian	IF1	6	1.276	0.049	1.444	0.046	2.720	46.912	Syllable
Italian	IF1	8	1.787	0.076	1.659	0.043	3.446	51.857	Syllable
Italian	IF1	9	1.4	0.036	1.722	0.056	3.122	44.843	Syllable
Italian	IF1	10	1.501	0.042	1.445	0.057	2.946	50.950	Syllable
Italian	IF1	11	1.516	0.039	2.072	0.072	3.588	42.252	Syllable
Italian	IF1	12	1.629	0.043	1.604	0.060	3.233	50.387	Syllable
Italian	IF1	13	1.588	0.037	1.506	0.042	3.094	51.325	Syllable
Italian	IF1	17	2.038	0.046	2.025	0.044	4.063	50.160	Syllable
Italian	IF1	20	1.73	0.044	1.831	0.061	3.561	48.582	Syllable
Italian	IF2	1	1.294	0.049	1.579	0.049	2.873	45.040	Syllable
Italian	IF2	3	1.328	0.042	1.651	0.058	2.979	44.579	Syllable
Italian	IF2	4	1.667	0.053	1.469	0.049	3.136	53.157	Syllable
Italian	IF2	6	1.375	0.050	1.579	0.056	2.954	46.547	Syllable
Italian	IF2	8	1.512	0.048	1.469	0.034	2.981	50.721	Syllable
Italian	IF2	11	1.626	0.053	2.028	0.058	3.654	44.499	Syllable
Italian	IF2	13	1.804	0.041	1.361	0.040	3.165	56.998	Syllable
Italian	IF2	15	1.537	0.028	1.903	0.056	3.440	44.680	Syllable
Italian	IF2	17	1.805	0.026	2.243	0.044	4.048	44.590	Syllable
Italian	IF2	18	1.745	0.043	2.015	0.053	3.760	46.410	Syllable
Italian	IF2	19	1.642	0.037	2.295	0.052	3.937	41.707	Syllable
Italian	IF2	20	1.687	0.041	2.096	0.049	3.783	44.594	Syllable
Japanese	JF1	1	1.586	0.044	1.341	0.039	2.927	54.185	Mora
Japanese	JF1	2	1.032	0.030	1.191	0.059	2.223	46.424	Mora
Japanese	JF1	3	1.179	0.039	1.112	0.042	2.291	51.462	Mora

			V-Intervals (s.ms)		C-Intervals (s.ms)		Total for Sentence		
Language	Speaker	Sentence[3]	Duration	Std. Dev.	Duration	Std. Dev.	Duration (s.ms)	%V	Rhythm Class
Japanese	JF1	5	1.539	0.059	1.071	0.039	2.610	58.966	Mora
Japanese	JF1	6	1.43	0.057	1.549	0.043	2.979	48.003	Mora
Japanese	JF1	7	1.518	0.043	1.446	0.047	2.964	51.215	Mora
Japanese	JF1	8	1.283	0.040	1.22	0.040	2.503	51.258	Mora
Japanese	JF1	9	1.243	0.028	1.439	0.053	2.682	46.346	Mora
Japanese	JF1	10	1.353	0.031	1.547	0.058	2.900	46.655	Mora
Japanese	JF1	11	1.517	0.035	1.502	0.043	3.019	50.248	Mora
Japanese	JF2	1	1.218	0.036	1.216	0.050	2.434	50.041	Mora
Japanese	JF2	2	1.077	0.036	0.922	0.029	1.999	53.877	Mora
Japanese	JF2	3	1.086	0.044	0.961	0.032	2.047	53.053	Mora
Japanese	JF2	4	1.4	0.057	1.132	0.033	2.532	55.292	Mora
Japanese	JF2	5	1.319	0.056	1.018	0.034	2.337	56.440	Mora
Japanese	JF2	6	1.422	0.064	1.392	0.046	2.814	50.533	Mora
Japanese	JF2	7	1.259	0.036	1.02	0.028	2.279	55.244	Mora
Japanese	JF2	8	1.052	0.032	1.094	0.042	2.146	49.021	Mora
Japanese	JF2	11	1.445	0.037	1.346	0.036	2.791	51.774	Mora
Japanese	JF2	12	1.417	0.031	1.245	0.037	2.662	53.231	Mora
Japanese	JM1	1	1.372	0.036	1.65	0.068	3.022	45.400	Mora
Japanese	JM1	3	1.09	0.043	1.32	0.046	2.410	45.228	Mora
Japanese	JM1	4	1.234	0.048	1.585	0.061	2.819	43.774	Mora
Japanese	JM1	5	1.536	0.051	1.39	0.044	2.926	52.495	Mora
Japanese	JM1	6	1.44	0.053	1.841	0.053	3.281	43.889	Mora
Japanese	JM1	7	1.465	0.040	1.321	0.036	2.786	52.584	Mora
Japanese	JM1	8	1.291	0.043	1.367	0.049	2.658	48.570	Mora
Japanese	JM1	10	1.299	0.045	1.227	0.037	2.526	51.425	Mora
Japanese	JM1	11	1.274	0.029	1.81	0.048	3.084	41.310	Mora
Japanese	JM1	12	1.206	0.023	1.589	0.060	2.795	43.148	Mora