Segment durations in a syllable frame

W. N. Campbell & S. D. Isard

Edinburgh University,
Centre for Speech Technology Research

80 South Bridge, Edinburgh, EH1 1HN UK.

running title: Segment durations in a syllable frame

Abstract

In continuous speech there is considerable variation in the durations of the segments. A measure of this variation in terms of elasticity of segments is proposed, and experiments are described that test the assumption that within a given syllable a constant factor of lengthening or compression can be applied uniformly to each segment in terms of its elasticity. Elasticity of segments is calculated on the basis of measured durations from a phonetically balanced 200-sentence database.

Comparisons were performed of relative compression and expansion of different segments with regard to position in the syllable and in the utterance. It was found that whereas segments in prepausal sentence-final syllables undergo greater lengthening in the ryme than in the onset, segments in sentence-internal syllables are lengthened or compressed more uniformly across the syllable. These findings offer a simpler account of the differential lengthening of vowels and consonants by explaining such differences in terms of elasticity about a mean.

We describe an implementation of the concept of segmental elasticity in the timning component of a text-to-speech system. In this system, duration is first computed at the syllable-level, and the segmental durations are accommodated to the syllable framework.

1 Introduction

Duration in speech is often analysed just in terms of the segmental composition of an utterance (Klatt 1974, Umeda 1988a,b, Crystal & House 1988) whereby inherent durations of each phoneme are modified by processes, operating at the level of the segment, that are sensitive to phonetic context and phrasal position. Much of the variance in duration can indeed be accounted for in this way, but if phonetic content were the only constraint on timing, then any higher-level rhythmic regularity in speech would have to be achieved through systematic modification of segment durations by rules insensitive to rhythmic constraints per se.

Many researchers (see for example Huggins 1982, Fowler 1977, Lehiste 1977, Scott 1982) have found that the interval between onsets of stressed vowels is perceptually significant. Carlson, Granström and Klatt (1979) show that differences in the interstress interval affect both the naturalness and the intelligibility of synthetic speech and suggest (ibid p.240) that the rules in the Klatt (1979) duration algorithm need to be improved to better model the interstress intervals found in natural speech.

In order to accommodate such rhythmic effects, we have implemented a syllable-based duration model for the text-to-speech system being developed at CSTR. Syllable durations are computed taking into account rhythmic and phrasal factors, such as stress and position in the phrase and foot, and individual segment durations are then accommodated to the syllable framework. Models for both syllable duration and segment accommodation are based on analysis of speech corpora.

Witten (1977) reported a previous algorithm based on accommodation of segment durations in predetermined foot timings but these were not based on statistics drawn from observations of speech data and failed to take into account other factors known to influence the segment durations.

Bartkova & Sorin (1987) describe a set of timing rules for speech synthesis in French, based on analysis of corpora. Their model differs from the one described here in that it is based on modifications to segment durations due primarily to phonetic effects rather than incorporating such effects into a higher-level framework.

2 Models of syllable and segmental timing

Two different corpora were used for observation of syllable and segment-level timing constraints. This was motivated by the need for prosodic naturalness sustained over periods of speech at the level of the paragraph and above in the one case, and the need for dense and phonetically balanced segment data in the other. As the specially constructed sentences of our phonetic database were read with the sole purpose of providing a source of speech data, and with no other communicative intent, it is not expected that they will be a reliable source of prosodic information.

2.1 Syllable timing

Campbell (1990a) has shown that an implementation of the Klatt rules for segment duration, adapted for British English and optimised for the test text, is capable of accounting for 70% of the variance observed in syllable durations of a broadcast reading of a short story. The text used was a four-thousand syllable passage from the Spoken English Corpus (Knowles and Taylor 1988), prosodically transcribed and measured for duration at the level of the syllable. No attempt was made to model any variation in duration arising from changes in the rate of presentation of the different sections of the passage.

Klatt (1974 p. 62), expressed agreement with "the contention that the syllable is an articulatory programming unit", but he went no further in the development of this in his model. Campbell (1990b) showed that a syllable-based model with minimal sensitivity to segment-specific information can account for as much of the durational variance at the syllable level as one based on segmental data. The model takes the following factors into account and predicts a duration for each syllable based on observations of corresponding durations in the corpus.

1. the number of phonemes in a broad transcription of the syllable.
2. the nature of the syllabic peak; whether it is a tense or lax vowel, a diphthong, or a sonorant consonant.
3. the position of the syllable in the foot.
4. the position of the syllable in the phrase and clause.
5. the stress assigned to the syllable. and the nature of any pitch movement associated with the syllable.
6. the function/content role of its parent word.

Rather than attempt to model the interactions and effects of these factors directly, a neural net was trained on the first thousand syllables of the corpus and tested with three subsequent sets of a thousand syllables each.

2.2 Segmental timing

2.2.1 Data

Two hundred phonetically balanced sentences, recorded and transcribed at CSTR as part of the SCRIBE database project were used as the source of segment durations. These sentences, read once in isolated word form and once as complete continuous sentences by an adult male speaker of RP English, cover most of the permissible demi-syllables in English with almost all combinations of vowels and single consonants (in both initial and final position) as well as providing examples of consonant clusters up to length four.

A team of phoneticians produced narrow phonetic transcriptions of the sentences according to the criteria in Laver et al (1989). These transcriptions provided the material from which our segment durations were obtained. No syllable boundary information is marked in the SCRIBE transcriptions, so it was necessary to group related segments first into words and then into syllables.

Segments in monosyllables were tagged as belonging to either onset, peak or coda, regardless of finer ordering distinctions within clusters of consonants or vowels, and an additional distinction was made for ambisyllabic segments in the case of polysyllables. All consonants preceding the first vowel or syllabic consonant after a word boundary were tagged as *onset*, and all following the last vowel and before a word boundary were tagged as *coda*. All vowels and syllabic consonants were tagged as *peak*.

Similar tagging was performed in the case of polysyllabic words, but any internal consonants were considered potentially ambisyllabic and tagged as *medial*. Syllabification was performed according to the following rules: A single medial consonant was assumed to be preceded by a syllable boundary and functioning as *onset*. A pair of medial consonants were assumed to have a syllable boundary between them, the first being *coda* and the second *onset*. A cluster of three or more medial consonants were assumed to have a syllable boundary between the second and the third.

An examination of the syllables produced in this way revealed no obvious cases of mis-syllabification and in all cases, the *medial* tagging was retained when a consonant was assigned to either onset or coda position so syllables created in this way were thereby distinguished from monosyllabic words.

Means (μ) and variances (σ^2) were calculated for each phoneme, and the individual segment durations then normalised by z-transform as in formula (1) below to a value that reflects the amount of variation in terms of standard deviations about each segment mean.

$$z_{token} = (observed\,duration_{token} - \mu_{type})/\sigma_{type} \tag{1}$$

z can be positive or negative, and for normally distributed data falls with a probability of 0.998 within the range of plus or minus 3. There is a positive skew in the SCRIBE segment data so that positive values extend further than negative ones; but this difference should have no important effect on the results of the experiments below, which involve comparisons only between similar extremes of the distribution.

Segments shorter than the mean will have a negative value of z, and those longer than the mean will have a positive value. In the following sections of the paper we refer to sements with positive z as *lengthened* and those with a negative z as *shortened.* This usage is merely a convenient shorthand and is not meant to suggest that the mean value represents natural, inherent or default durations.

2.2.2 The elasticity hypothesis

The strong form of the elasticity hypothesis (Campbell, in preparation) says that all segments in a given syllable fall at the same place in their respective distributions. That is, for any given syllable, there should be a number k of standard deviations such that the length of every segment in the syllable is equal to $\mu_{seg} + k\sigma_{seg}$, where μ_{seg} and σ_{seg} are the mean and standard deviation respectively of durations of the particular segment type.

The z value determined above, a quantification of each segment's position in its distribution, can be taken as a measure of its state relative to its elasticity. A segment with a high variance (such as a tense vowel) that shows considerable difference in duration in terms of absolute millisecond measurements may be in the same relative state of expansion or compression as one with a much smaller variance (a stop, for example) that appears to change less in absolute terms.

By way of an example, consider the words *at* and *add.* In our segment database, the /a/ has a mean duration of 119 ms (sd = 37 ms), the /t/ a mean of 41 ms (sd = 21 ms), and the /d/ a mean of 39 ms (sd = 19 ms). The mean durations of the two words, all else being equal, would therefore sum to 160 ms and 158 ms respectively. In a context where the total duration of *at* was in fact longer, say 300 ms, both segments would be lengthened in accordance with their distributions by a constant number k of standard deviations so as to make $(\mu_a + k\sigma_a) + (\mu_t + k\sigma_t) = 300$. Solving for k and substituting the result back into the formula, we find that the duration of the /a/ in *at* would become 208 ms, and the /t/ 92 ms. For in *add*, the same exercise yields an /a/ duration of 213 ms, and a /d/ of 87 ms. The overall word lengths are the same, but because the elasticity of the /d/ is less than that of the /t/, the vowel lengthens, by 5 ms in this case, to accommodate. In neither additive nor multiplicative terms is there a linear relationship between the amount of change undergone by each segment to fit the new framework; each has increased by a little more than 2.4 standard deviations.

Weaker forms of the elasticity hypothesis would state that statistics have to be gathered separately for syllables in different positions in the sentence (e.g. finally versus non-finally), for segments in different parts of the syllable (e.g for those in the onset and ryme), and in different phonetic contexts (e.g for vowels before voiced and unvoiced stops).

3 Experiments

Several experiments were performed to determine the extent to which the state of each segment, measured in terms of its elasticity, is uniform throughout the syllable. We are testing the strong assumption that lengthening applies equally across all components in a syllable. In particular we are interested in whether there is a distinction between consonants and vowels, where under millisecond observations most difference would be likely to be found. The elasticity hypothesis predicts that the z-scores for each segment will be the same within each syllable. We looked at sentence-final syllables separately from non-final syllables, and at long, short and intermediate categories in our analysis.

3.1 Long and short syllables

The raw duration of a syllable in units of time is no measure of its relative length, being the result of so many interacting factors, and can not be used as a reliable measure of lengthening as defined here. The normalised values for each segment within a syllable were therefore averaged and the result taken to be representative of the length of the syllable as a whole. The one sd cutoff ($z > 1$ or $z < -1$) was taken as criterial, and all syllables whose averaged value was greater than +1 were taken as *long*, and those whose averaged value was less than −1 were considered *short*. Syllables in sentence-final position were excluded from this grouping as the effect of pre-pausal lengthening is being considered separately.

The distribution of the values of the tokens in both groups was tested to ensure that the individual scores of the component phonemes were representative, and in neither case did more than a few outliers exhibit a different sign from the rest of the group. Because of the tendency to maximise onset in the syllabification procedure, medial segments showed a clear bias in their distribution. Of the 63 medial segments in the *long* group, only 15 are in coda position, and the remaining 48 in the onset.

The means and standard deviations of the values in each group were determined and the overall mean found to be 1.4 (sd 1.07; n = 439) for the *long* group, and −1.2 (sd 0.54; n = 344) for the *short* group. Table I shows the individual means for each segment position within the three syllable groups.

TABLE I HERE

3.2 Discussion

In the *short* group, the compression appears to be constant across all syllable parts. The small differences in the means of the peak and coda segments were not significant (t = 1.126; ns; df = 205), and it would appear from these results that a single factor of compression can be used to describe the shortening undergone by the segments in these syllables. Nor is there any significance in the difference between means for onset and peak segments in the *long* group (t = 0.67; ns; df = 287).

The difference in means of onset and coda consonants in the *long* group, however, is significant (t = 3.62; $p < 0.001$; df = 187), as is the difference in means for peak and coda segments (t = 2.98; $p < 0.01$; df = 272). A similar significant shortening (t = 3.09; $p < 0.01$; df = 3501), of coda segments is also found in the reference group of *intermediate* syllables whose values fall between +1 and −1.

This differential lengthening was initially thought to be a consequence of the different articulatory environment of the two consonant groups, perhaps requiring a dark/light distinction to be made for all consonants, but a test with data normalised for onset and coda segments separately revealed similar differences.

4 Final-lengthening

Many researchers (see for example House & Crystal 1988) have noted the lengthening applied to word-final segments in a pre-pausal position. Since our database does not include sentence-internal pauses, consisting of single sentences constructed with the purpose of maximising segment-segment coarticulations, the only pre-pausal contexts we have are those in sentence-final position.

It is likely that there is also phrase-final lengthening (Klatt 1979) in our data but the absence of phrase-level annotation of the sentences means that this cannot be studied separately here.

Since word-final position is also by definition syllable-final, the first test was to determine whether and to what extent the higher-level label was required. Similarly, further factoring was performed to determine whether lengthening is uniform across the syllable. To this end, means were determined for segments in sentence-final words, in sentence-final syllables, in absolute sentence-final position, and within the latter group, to vowels in absolute sentence-final position. Table II shows results for these factorisations.

TABLE II HERE

The difference in means between segments in sentence-final words and sentence-final syllables is significant (t = 3.89; $p < 0.001$; df = 1876), as is that between segments in sentence-final syllables and in absolute sentence-final position (t = 4.096; $p < 0.001$; df = 966). Because the mean for final syllables is higher than that for final words, we assume that the bulk of final-word lengthening is actually taking place in the final syllables, and confine our attention to lengthening as it is found there.

Sentence-final segments are longer than non-final segments in the final syllables, and we have examined the distribution of lengthening within the syllable parts. Means and standard deviations were compared for the onset peak and coda segments of sentence-final syllables. Table III shows values for segments in this group factored by position in onset, peak and coda. By definition, all *medial* segments in sentence-final syllables are in onset position.

TABLE III HERE

The difference between means for onset and peak segments is significant (t = 7.08; $p < 0.001$; df = 392), but no significance was found in the difference between peak and coda at (t = 0.448; ns: df 485). The mean for the subgroup of vowels in sentence-final position is significantly higher than that of the peak subgroup as a whole, in which they are included, and on which they presumably have an inflationary effect (t = 4.26; $p < 0.001$; df = 441). This is not simply attributable to a lengthening of vowels in open syllables since the value for the group of word-final vowels as a whole is actually less than zero (z = -0.14; sd = 0.91; n = 811).

4.1 Discussion

Means become larger as the unit size (word, syllable or segment) becomes smaller, and there is a significant difference between the lengthening measured in onset segments in pre-pausal syllables and in those occurring in the peak and coda. It would appear from this that lengthening in pre-pausal syllables can be interpreted as a boundary phenomenon, confined to the ryme, and therefore different from the lengthening observed in sentence-internal syllables which appears to be a property of the syllable as a whole, and which is smaller in that case for the coda than for the rest of the syllable.

5 Phonetically-motivated lengthening

Lengthening of vowels before voiced consonants often has been discussed in the literature (House & Fairbanks 1953, Peterson & Lehiste 1960, Luce & Charles-Luce 1985, and Crystal & House 1988). In section 2.2.2 above, we showed how a syllabic peak can lengthen by accommodation alone to fit in with the typically shorter voiced stop.

If non-pre-pausal length, in terms of elasticity, is constant across the syllable, and applied without respect to phonetic features, then the strong form of the elasticity theory would predict that there should be no further differential effect such as might help cue the voicing distinction of the following stop.

Table IV shows results of comparisons made first for differential lengthening within all syllables, then in syllables factored for final and non-final positions, in long and short syllables, and finally for an equivalent effect in open syllables which were followed by a stop onset to the following word, across the word boundaries.

TABLE IV HERE

Table V gives a comparison of vowels preceding voiced fricatives.

TABLE V HERE

In neither the general case nor the separately factored cases of lengthened and shortened variants was the difference in the lengthening of vowels before voiced and voiceless fricatives found to be significant. Although vowels are found to be lengthened more before voiced fricatives than before unvoiced ones, the effect is not significant, so we have not found strong evidence for a general phonetic effect cued by the feature of voicing alone. There is at the very least an interaction with manner of articulation.

5.1 Discussion

In all cases, the normalised values for vowels before voiced stops were greater than those for vowels before voiceless stops. The differences are significant for the groups of all syllables, and all sentence-internal syllables, but not for smaller groups.

The difference appears to be large in sentence-final syllables and in this case the lack of statistical significance in the result may be due to the small number of tokens of voiced phrase-final stops in the data.

Very little difference in vowel length was found when the vowel and following stop were separated by a word boundary, so that lengthening before voiced stops appears to be a property of the syllable rather than a general phonetic effect that occurs at all times in continuous speech.

6 Timing in the CSTR text-to-speech system

6.1 Syllable duration

In the timing component of the text-to-speech system that we are developing, syllable duration is computed according to the model of section 2.1 above. The model is implemented as a 3-layer neural net, which is trained by back-propagation (McLelland, Rumelhart et al 1986) on the log-transformed durations from the syllable database. The effect of the log-transform is to bring the syllable durations close to a normal distribution in order to make it easier for the network to model the variation. Given a 6-tuple of features the net returns a syllable duration in log milliseconds.

6.2 Segment duration

After the overall duration for a syllable has been determined, individual segment durations are accommodated using a modified form of the elasticity hypothesis. The first modification is to reduce the positive skew in the data by transforming segment durations to log milliseconds. Applied to the log-transformed data, the strong form of the elasticity hypothesis would predict that for any given syllable, there should be a factor k such that each segment token belonging to the syllable would have duration $exp(\mu + k\sigma)$, where μ and σ are the mean and standard deviation respectively of log-transformed durations corresponding to the segment type. For non-final syllables, this is in fact the way in which segment durations are assigned. The factor k is determined by solving the equation

$$\Delta = \sum_{i=1}^{n} exp(\mu_i + k\sigma_i) \tag{2}$$

for k, where:
Δ is the duration determined for the whole syllable,
n is the number of segments in the syllable,
and μ_i is the mean and σ_i the standard deviation of the log transform of measurements for all the tokens in the database corresponding to segment_i.

For final syllables, k is determined as above, but the ith segment of the syllable is assigned a duration of $exp(\mu_i + 0.75^{(n-i)}k\sigma_i)$. This means a) that segments at the end of the syllable are lengthened more than earlier ones and b) that the duration of final syllables is shorter than the syllable duration model dictates. a) represents an approximation to our findings in section 4 above which needs to be further refined. b) seems to be fairly satisfactory on an empirical basis, in that although the syllable duration model was predicting the lengths of final syllables accurately in terms of our data, final syllables sounded too long in our synthetic speech. It is possible that this perceptual effect arises from the fact that we do not yet control for decay of amplitude into the pause.

Differential vocalic lengthening before stops is modelled only in so far as it can be accounted for by the greater elasticity of unvoiced stops, as in the *at* vs. *add* example of section 2 above.

Manipulation of the parameter k gives us some control of articulation rate in our synthetic speech. Adding the same fixed value to the computed k for every segment preserves the rhythmic characteristics of the utterance while providing a convincing rendering of fast and slow articulation. Further work on the modelling of pause durations is required for a complete model of speech rate.

7 Conclusion

We have presented a model of segmental duration, implemented in a speech synthesis system, in which durations are first computed at the syllable level, taking rhythmic factors into account and ignoring fine phonetic detail. Segment durations are accommodated to the syllable framework in a simple way, using a measure of segment elasticity.

The elasticity hypothesis on which the model is based was tested in a number of experiments. The results of these experiments showed that the strong form of the hypothesis applied better to non-final syllables than to final ones. Even within non-final syllables, however, the hypothesis needs revision to take full account of phonetic effects such as differential vowel lengthening before voiced and unvoiced stops.

These results suggest that timing processes may operate at three different levels: at the level of the phrase, where it may be related to boundary phenomena, at the level of the syllable, as a result of stress and rhythmic effects, and at the level of the segment to accommodate such effects as the increase in length that is associated with a following voiced stop. A similar conclusion was reached by Edwards & Beckman (1988) with data from articulatory measures of syllable timing.

In our current speech synthesis system we model phrase and syllable level effects, but not yet intra-syllable phonetic effects.

References

Bartkova K. & Sorin C. (1987) A model of segmental duration for speech synthesis in French. *Speech Communication* **6** 245 - 260.

Campbell W. N. (1990a) Measuring speech-rate in the Spoken English Corpus, in *Theory and Practice in Corpus Linguistics*, (J. AArts & W. Meijs, editors) pp 61 - 81. Rodopi, Amsterdam.

Campbell W. N. (1990b) Analog i/o nets for syllable timing in *Speech Communication Special Issue on Neural Nets and Speech*, **9** (1). Elsevier Science Publishers B. V. (North Holland).

Campbell W. N. (in preparation) PhD Thesis, University of Sussex.

Carlson, Granström & Klatt (1987) Some notes on the perception of temporal paterns in speech in *Frontiers of Speech Communication Research*, (Lindblom & Ohman, editors), pp 233 - 243. Academic Press.

Crystal T. H. & House A. (1988) Segmental durations in connected speech signals: Current results. *Journal of the Acoustical Society of America* **83**(4) 1553 - 1573.

Edwards J. & Beckman M. (1988) Articulatory timing and the prosodic interpretation of syllable duration, *Phonetica* **45**,156 - 174.

Fowler C. A. (1977) Timing control in speech production. Indiana University Linguistics Club, Bloomington.

House A. S. & Crystal H. (1988) Segmental durations in conected speech signals: Syllabic stress, *Journal of the Acoustical Society of America* **83** 1574 - 1585.

House A. & Fairbanks G. (1953) The influence of consonant environment upon the secondary acoustical characteristics of vowels *Journal of the Acoustical Society of America* **25** 105 - 113.

Huggins A W. F. (1972) On the perception of temporal phenomena in speech, *Journal of the Acoustical Society of America* **51** (4) 1279 - 1290.

Knowles G. & Taylor L. (1988) *Manual to accompany the Spoken English Corpus*, IBM UK Scientific Centre & Lancaster University.

Klatt D. H. (1974) The duration of [S] in English words, *Journal of Speech and Hearing Research*, **17** 51 - 53.

Klatt D. H. (1976) Linguistic uses of segment duration in English. *Journal of the Acoustical Society of America* **59** 1208 - 1221.

Klatt D. H. (1979) Synthesis by rule of segmental durations in English sentences. In *Frontiers of Speech Communication Research* (Lindblom & Ohman, editors), pp 287 - 300. Academic Press.

Klatt D. H. (1987) Review of text-to-speech conversion for English. *Journal of the Acoustical Society of America* **82** 737 - 693.

Laver J. Hieronymus J. Dalby J. Hiller S. McAllister M. Purves E. M. (1989) *ATR/CSTR Speech Database Project Status Report* No.3, Centre for Speech Technology Research, University of Edinburgh.

Lehiste I. (1977) Isochrony reconsidered, *Journal of Phonetics* **5**, 253 - 265.

Luce P. A. & Charles-Luce J. (1960) Contextual effects on vowel duration, and the consonant/vowel ratio in speech production. *Journal of the Acoustical Society of America* **78** 1949 - 1957.

McLelland J. L., Rumelhart D. E. and the PDP Research Group (1986) *Parallel Distributed Processing: Explorations in the microstructure of cognition*, MIT Press, Cambridge MA.

Peterson G. E. & Lehiste I. (1960) Duration of syllable nuclei in English *Journal of the Acoustical Society of America* **32** 693 - 703.

Scott D. R. (1980) Duration as a clue to the perception of a phrase boundary. *Journal of the Acoustical Society of America* **71** 996 - 1007.

Umeda N. (1988a) Vowel duration in American English. *Journal of the Acoustical Society of America* **58** 434 - 445.

Umeda N. (1988b) Consonant duration in American English. *Journal of the Acoustical Society of America* **61** 846 - 858.

Witten I. H. (1977) A flexible scheme for assigning timing and pitch to synthetic speech, *Language & Speech* **20** 240 - 260.

Acknowledgement

This work was supported by the Information Engineering Directorate/Science and Engineering Research Council as part of the IED/SERC Large Scale Integrated Speech Technology Demonstrator Project (SERC grants D/29604, D/29628, F/10309, F/10316, F/70471) in collaboration with Marconi Speech and Information Systems and Loughborough University of Technology.

Table I: Z values for three groups of syllables.

	long syllables:			short syllables:			intermediate syllables:		
	mean	sd	n	mean	sd	n	mean	sd	n
onset	1.56	0.93	102	-1.22	0.56	99	-0.05	0.89	1793
peak	1.47	1.16	187	-1.22	0.51	170	-0.09	0.79	2881
coda	1.03	1.08	87	-1.12	0.38	37	-0.14	0.83	1710
medial	1.48	0.92	63	-1.26	0.55	37	-0.08	0.88	1042

Table II: Values for segments in different sentence-final contexts.

	Words:	syllables:	segments:	vowels:
mean	0.60	0.82	1.34	1.63
sd	1.23	1.32	1.44	1.42
n	1154	770	200	49

Table III: Values for segments by position in final syllables.

	mean	sd	n
onset	0.33	0.97	232
peak	1.09	1.25	245
coda	1.14	1.21	242

Table IV: Values for vowels followed by voiced and unvoiced stops.

	unvoiced:			voiced:			significance:		
	mean	sd	n	mean	sd	n	t	df	p
all syllables	-0.022	0.89	374	0.227	0.98	203	3.09	575	<0.01
sent-internal sylls	-0.105	0.83	345	0.125	0.84	190	3.05	533	<0.01
sent-final sylls	0.959	1.04	29	1.714	1.53	13	1.87	40	n.s.
long syllables	0.703	0.65	176	0.844	0.84	114	1.59	288	n.s.
short syllables	-0.667	0.48	198	-0.564	0.39	89	1.78	285	n.s.
across word boundary	-0.313	0.86	125	-0.308	0.86	112	0.29	235	n.s.

Table V: Values for vowels before voiced and unvoiced fricatives:

	unvoiced:			voiced:			significance:		
	mean	**sd**	**n**	**mean**	**sd**	**n**	**t**	**df**	**p**
all syllables	0.132	0.965	298	0.248	1.107	388	1.44	684	n.s.