

# RHYTHM AND VOWEL QUALITY IN ACCENTS OF ENGLISH

**GEOFF SCHWARTZ**

geoff@ifa.amu.edu.pl

Adam Mickiewicz University

## **Abstract**

This paper will examine rhythmic differences among native and non-native accents of English, and report on a pilot experiment investigating a hypothesized interaction between rhythm and vowel quality. A new metric, % SteadyState, an acoustic measure that quantifies the purity of vowels, appears to capture rhythmic differences that have been reported among various native and non-native accents of English. In the tradition of other recently developed rhythm metrics, these findings suggest a link between rhythm and segmental phonology. Additionally, the perspective gained from this study may be beneficial to learners whose goal is native-like vowel quality, offering an understanding of the dynamic properties of English vowels.

**Keywords:** phonology, rhythm, vowel quality, English pronunciation

## **1. Introduction**

Considering the geographic and sociolinguistic diversity among speakers of English, it is no surprise that we find significant variation in all aspects of the linguistic system, from vocabulary and syntax to phonology. In the area of pronunciation, different accents of English are frequently described in terms of variation in segmental phonology, particularly in the vowel system. Additionally, variation has been documented in the realization of suprasegmental features such as intonation and rhythm. Rhythmically English is traditionally classified as a stress-timed language. However, the great diversity of accents with which English is spoken suggests the possibility of rhythmic variation, which has indeed been found. Stress-timing is dominant in RP and other accents based in the South of England, as well as in the United States and Australia. Varying degrees of syllable timing have been reported in Singapore (Low et al 2000), Wales (White and Mattys 2007), and Scotland (White and Mattys 2007).

One question that seems not to have received significant attention in the study of English accents is the possibility of an interaction between segmental features such as the vowel system, and suprasegmental rhythm. Although traditionally seen as independent areas, there are reasons to believe that segmental phonology and rhythm must be linked. For instance, reduction typically affects vowel quality as well as duration. Additionally, diphthongization typically is more robust on long vowels, and the presence of long vowels in a language is connected with durational variability that is a marker of rhythm. This paper will present a preliminary acoustic study of various

accents of English. Our goal is to explore the hypothesis that vowel quality may reflect rhythmic properties of a language.

In my experience teaching American English pronunciation to Polish learners, I have noticed two basic features of English phonology that present the most serious obstacles to the acquisition of native-like pronunciation. The first of these is the vowel system. In particular, the dynamic properties of American vowels (Strange 1989) cause problems for speakers of languages like Polish whose vowels are very pure in quality. The other issue is vowel reduction. Since unstressed vowels in Polish undergo minimal vowel reduction, L1 interference frequently causes errors in the pronunciation of unstressed English vowels. These two difficulties encountered by Polish learners would seem at first glance to be independent. One concerns vowel quality, a segmental feature, while the other affects suprasegmental rhythm. However, if we adopt a holistic view of linguistic structure (Donegan and Stampe 1983), we must immediately raise the question of whether and how these two areas interact.

The rest of the paper will proceed as follows. Section 2 will discuss the role of phonological considerations in the classification of linguistic rhythm. Section 3 will discuss perceptual underpinnings behind the rhythm-vowel quality connection. Section 4 will introduce vowel quality variation of across different accents of English. Section 5 presents the experiment. Section 6 concludes with some general discussion.

## 2. The Phonology of rhythm

The rhythm class hypothesis goes back at least as far as Pike (1945), who built on the characterizations of “machine gun” rhythm for Italian and Spanish and “Morse code” rhythm for English and Dutch. This work spawned a hypothesis of isochrony in speech (Abercrombie 1967). In stress-timed languages isochrony between stresses was assumed, while syllable-based isochrony was assumed for syllable-timed languages. Experimental studies, however, have failed to find solid evidence of isochrony in the speech signal, suggesting that rhythm is a perceptual, rather than acoustic phenomenon.

Recent work has therefore concentrated on phonological criteria that have traditionally been linked to the rhythm class hypothesis (e.g. Donegan and Stampe 1983, Bertinetto 1989). Two such phonological observations have been employed in an attempt to quantify the assumed rhythmic typology. Vowel reduction in stress-timed languages is assumed to result in greater durational variability, measured using metrics calculated from the durations of vocalic intervals. These reduction-based metrics include Pairwise Variability Indices (PVI; Grabe and Low 2002), DeltaV (Ramus et al 1999), the standard deviation of vocalic intervals, and VarcoV (Dellwo 2006), the standard deviation of vocalic intervals divided by the average vowel duration. The other phonological criterion employed in rhythm metrics is phonotactics. Phonotactic measures include %V and %C (Ramus et al 1999), representing the total percentages of vocalic and consonantal intervals, respectively. The underlying hypothesis behind these measures is that the codas and consonant clusters found in stress-timed languages contribute to a greater consonantal portion of the signal.

These phonological rhythm metrics have been employed with varying degrees of success in describing how rhythm type might be reflected in the speech signal. In a study

attempting to evaluate rhythm metrics in terms of assumed distinctions in rhythmic typology, White and Mattys (2007) found that rate-normalised metrics based on vowels, VarcoV and nPVI-V, were the most reliable in capturing the hypothesized rhythm types. This finding suggests that of the two phonological criteria employed in the recently developed rhythm metrics, vowel reduction and variability in vowel duration is a better indicator of rhythm type than phonotactics.

The premise behind phonotactically-based rhythm metrics is that complex structures containing numerous consonants produce heavier “syllables” that attract stress in stress-timed languages. This claim may be called into question due to its reliance on the syllable as a unit of processing. Syllabic structures are language-specific perceptual constructs, as evidenced by the fact that the acoustic input /pi:k/ is heard as one syllable by English listeners and as two syllables by Japanese listeners. If rhythm is intended to be classified according to universal phonetic properties, language-specific syllable structures should not play a role. Additionally, the phonotactic approach to rhythm predicts that Polish, with its rich inventory of consonant clusters, should classify as stress-timed. Other phonological criteria, however, point to Polish as a syllable-timed language. These criteria include full vowels in unstressed positions, fixed stress, and a simple vowel inventory. Native speaker intuitions also point to Polish as being syllable-timed. These intuitions are supported by findings in Wagner (2007).

In light of the previous discussion, we may conclude that of the two phonological criteria employed in the development of rhythm metrics, vowel reduction offers more insightful measures of rhythm than phonotactics. The implication here is that the perception of rhythm stems largely from properties associated with the vocalic intervals in speech.

### 3. Motivating the link between rhythm and vowel quality

Under the assumption that rhythm is reflected in phonological behaviour, it is reasonable to expect that other phonological features, particularly those associated with vowels, may contribute to the perception of a language's rhythm. For example, Donegan and Stampe's (1983) holistic typology suggests a possible link between rhythm and vowel quality. They note that stress-timed languages typically have larger vowel inventories that may include diphthongs and diphthongised monophthongs, while non-stress languages have simpler systems and vowels that are relatively pure in quality. If these predictions are applied to the rhythm class hypothesis, we may consider the possibility that in addition to duration, spectral properties of vowels may provide additional insightful metrics for the description of linguistic rhythm. This paper will present a preliminary acoustic study exploring the hypothesis that rhythm may be reflected in vowel quality.

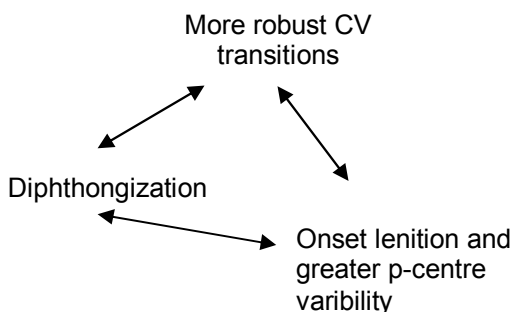
The motivation behind the premise that vowel quality may reflect rhythmic properties is rather intricate in nature, touching on a number of diverse phonological and perceptual issues. As a starting point, consider the following phonological observations. Syllable-timed languages such as Italian tend to have vowels that are more pure in quality than those found in the stress-timed languages such as English. Additionally, in

stress-timed languages such as English and Danish, we may observe a more significant amount of consonant lenition than in syllable-timed languages.

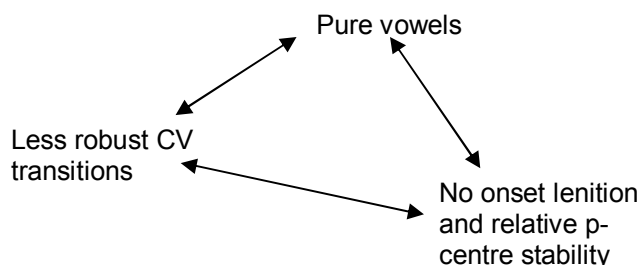
Although these observations are not obviously related, perceptual considerations may suggest a possible interaction. Briefly stated, the hypothesis goes as follows. Vowels frequently function as acoustic background to the linguistic message, transmitting information about neighbouring consonants (Wright et al 1997), a fact that may have two consequences for the phonology of a language. First, CV transitions may be incorporated into the canonical representations of vowels themselves, resulting in dynamic specification (Strange 1989), in which formant trajectories, rather than static targets, form the basis of listener identification. Additionally, robust CV transitions that allow listeners to reconstruct consonants provide a perceptual license for consonant lenition – weaker consonant articulations have minimal perceptual consequences when the consonant may be reconstructed on the basis of vowel formants. In other words, when lenition is tolerated we may assume that more of the vowel serves to cue listeners to the identity of neighbouring consonants, promoting the functional role of its dynamic properties.

The connection between lenition and rhythm assumed here is founded on experimental findings in which the manner of articulation of an onset consonant affects listener identification of the p-centre, the perceptual “moment of occurrence” of a syllable (e.g. Scott 1993). In languages that have a great deal of consonant lenition, we should expect greater variability in strength, and greater variability in p-centre location. When vowels are pure in quality we may hypothesize that the reverse is true: no lenition and lesser variability in onset strength and p-centre location. Given the language-specific nature of speech perception (e.g. Boersma and Hamann 2009), we may speculate that p-centre location may be subject to further, listener induced adjustments.

In sum, our rhythm hypothesis may be represented as a set of three phonological phenomena that are predicted to interact. These interactions are schematized in (1) (for stress-timing) and (2) (syllable-timing) below.



(1) Hypothesized interaction of vowel quality and stress-timing



(2) Hypothesized interaction of vowel quality and syllable-timing.

Notice that no causality is assumed in these schematics. Also worth noting is that an interaction is predicted between robustness of CV transitions and vowel quality. To my knowledge, such an interaction has not been the subject of systematic experimental study<sup>1</sup>, and could serve as the basis for a new and interesting research programme.

#### 4. Vowel quality in accents of English

In considering the broad diversity of accents of English, we may notice that variation is found largely in the vowel systems. The two standard accents that most frequently serve as models for the teaching of English, Received Pronunciation and General American, save for rhoticity<sup>2</sup>, are for the most part equivalent in their consonant systems. This fact may be assumed to stem from the functional role of vowels as acoustic background for the perception of neighbouring consonants, which allows for vocalic variation with minimal consequences for intelligibility. Jenkins (2000) exploits this idea in her *Lingua Franca Core* proposal, noting that vowel quality is rarely a cause of communication breakdown in conversations between non-native speakers of English.

Much of the vowel quality variation that we find across accents of English may be described in terms of the degree of diphthongization of stressed monophthongs, an idea that is reflected in the fact that different symbols are used for the transcription of the vowels in the words *sheet* and *shoot*. An informal survey of such vowels suggests that diphthongisation is more significant in those dialects of English that are said to be characterized by a greater degree of stress-timing. These include RP, General American, and Australian English. Conversely, diphthongisation is minimal in Scottish and Welsh accents that have been reported to have a more syllable-timed rhythm.

<sup>1</sup>Hume et al (1999) found that Korean listeners show somewhat greater sensitivity to formant transitions than American listeners. This finding says nothing about the relative robustness of the transitions in the two languages, since only one set of stimuli, produced by a Korean speaker, was used for the perception test. Hume et al speculate that Koreans' sensitivity to formant transitions for place identification might stem from the fact that this language has a three way laryngeal contrast, so burst noise and aspiration must serve to cue manner.

<sup>2</sup>Even rhoticity may be considered largely a vocalic phenomenon.

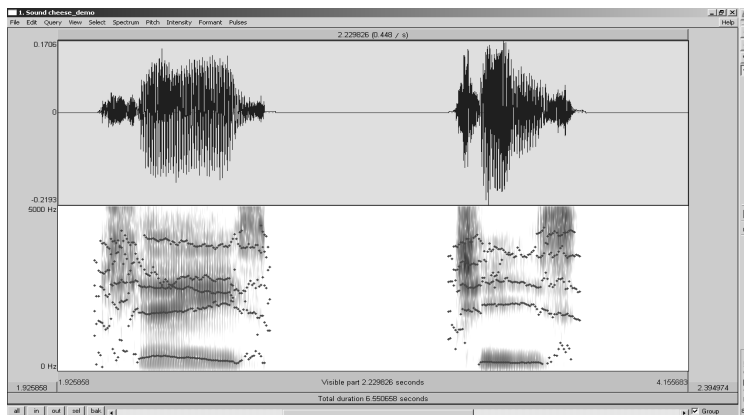


Figure 1 – Two tokens of the word *cheese*, Southern England (left) and Scotland.

For an illustration of this connection, consider Figure 1, which depicts a waveform and spectrogram of two tokens of the word *cheese*. The token on the left is produced by a Southern British English speaker, while the token on the right is produced by speaker from Scotland. Notice that the first and second formants on the spectrogram of the Scottish token stay more or less on the same level throughout the vowel, indicating pureness of vowel quality. The English token on the other hand shows diphthongisation, reflected in the display in the movement of the two formants. F1 and F2 at the beginning of the vowel are relatively close together, indicating a lowered, less fronted articulation. Later in the vowel the formants move apart towards a peripheral, /i/-like target.

Measurements of F1 and F2 taken at onset and offset of these vowels allow us to quantify the diphthongization. For the English token F1 ranges from 478Hz at onset to 350 Hz at offset, while F2 ranges from 1920 Hz at onset to 2160 Hz at offset. Converted into Bark, a psychoacoustic scale allowing better quantification of auditory sensitivity to changes in acoustic frequency, these measures translate to overall F1 movement of 1.1 Bark and F2 movement of 1.0 Bark. The Scottish token, on the other hand shows both F1 and F2 movement of 0.27 Bark and 0.26 Bark, respectively. Thus, on a perceptual scale, the formant movement associated with the English token is roughly four times as salient as that of the Scottish token. This connection is studied systematically in the experiment described in the next section.

## 5. Experimental method

This section will describe an acoustic study of 14 different accents of English. Acoustic measurements are made with the goal of investigating a possible correlation between vowel purity and variability in vowel durations.

## 5.1 Data

Fourteen samples from the Speech Accent Archive ([accent.gmu.edu](http://accent.gmu.edu)), in which speakers of different accents all read the same text, served as data for the experiment. The archive is particularly useful because in principle all speakers are attempting to produce identical targets, facilitating comparisons across speakers. The accent breakdown of the samples is summarised in Table 1.

Table 1 – Accents investigated in the acoustic study

Origin	Number of accents analysed
Southern England	3
Scotland	3
Wales	2
Ireland	2
Southern USA	1
Poland	2
<i>Total</i>	<i>14</i>

## 5.2 Acoustic measurements

Two basic types of acoustic measurements were made on the data. The first sought to quantify vowel durations, and the second vowel quality. All measurements were made by hand using the Praat speech analysis program with a time-aligned waveform and spectrogram display. Vowel durations were measured in milliseconds for all vowels in each sample. Vowel quality was investigated on stressed vowels in the samples. First and second formant frequencies (in Hertz, later converted to the Bark scale) were measured at vowel onset, midpoint, and offset. Finally the portion of each vowel marked by a spectral steady state was measured (in milliseconds).

## 5.3 Calculations made on acoustic measurements

Two measures were calculated in the basis of the durations of vocalic intervals. The first, VarcoV, represents the standard deviation of all the vowels in a sample, divided by the average vowel duration multiplied by 100 to yield a percentage. A rate-normalized Pairwise Variability Index, the mean of the differences between successive vowels divided by the sum of vowel durations, was also calculated for vowel durations in each sample.

Four measures of vowel quality were calculated. Total formant excursion in Bark was calculated for both F1 and F2. An additional measure took the average of the two formant excursion measures. Finally, the steady state portion of each vowel was divided

by overall duration to yield a measure of vowel purity. A summary of all the metrics calculated for these data is given in Table 2.

Table 2 – Metrics calculated from acoustic measurements

Metric	Description
VarcoV	Standard deviation of vowel durations divided by mean vowel duration (times 100).
PVI-V	Mean of absolute value of differences between successive vowel durations divided the by sum of those durations (times 100)
F1 excursion	Sum of absolute value of differences (in Bark) between F1 at onset and midpoint, and midpoint and offset
F2 excursion	Sum of absolute value of differences (in Bark) between F2 at onset and midpoint, and midpoint and offset
F1/F2 excursion	Mean of F1 and F2 excursion values
%Steady	Portion of each vowel characterized by spectral steady state divided by vowel duration

## 5.4 Results – mean values

The mean values for each of the calculated metrics for each accent are given in Table 3. They are arranged in ascending order for VarcoV.

Table 3 – Results of metrics for individual accents

Accent	Varco V	PVI-V	F1 exc	F2 exc	F1/F2 exc	%Steady
Poland (J)	35	46.1	1.45	0.6	1.03	62
Poland (Mi)	36	47.4	0.69	0.79	0.74	55
Wales (Wr)	37	43.6	0.84	0.75	0.8	51
Poland (Ot)	38	61.3	0.85	1.11	0.98	56
Wales (Ca)	39	55.8	1.02	0.99	1.01	53
Scotland (Gl2)	49	57.5	0.74	0.7	0.72	48
Scotland (Gl)	50	61.4	0.82	0.75	0.79	56
Scotland (Ed)	51	75.9	1.3	1.09	1.2	41
Ireland (Du)	52	70.2	0.92	1.3	1.11	50
Ireland (Be)	55	43.2	1.16	1.43	1.3	43
England (Lh)	55	43.1	2.13	1.22	1.68	49
England (Du)	57	65.1	1.06	1.18	1.12	43
England (St)	62	65.4	1.82	0.85	1.34	49
Usa (La)	63	77.8	2.13	1.76	1.95	42



## 5.4 Results – correlations and accent comparisons

Pearson  $r$  correlation coefficients were calculated between each of the vowel reduction metrics and each of the vowel quality metrics. These correlations are summarised in Table 4.

Table 4 – Pearson  $r$  correlation coefficients between pairs of metrics

VarcoV-%Steady	-0.74
VarcoV-F1	0.6
VarcoV-F2	0.61
VarcoV-F1/F2	0.7
PVI-%Steady	-0.72
PVI-F1	0.17
PVI-F2	0.41
PVI-F1/F2	0.3

The correlation tests show relatively high correlations between Varco V and each of the vowel quality parameters. PVI showed a strong correlation only with % SteadyState. Among the vowel quality parameters, % Steady State showed the strongest (inverse) correlations with vowel reduction measures.

Significant differences between accents were found in a number of accent comparisons in which accents were grouped according to origin. VarcoV was greater for the Southern English accents than for Scottish, Welsh, and Polish ( $p < 0.05$  for each comparison). This reflects the standard view that Southern British English represents a classic example of a stress-timed language. Steady state percentage was significantly greater for the Polish accents as a group than the southern English accents ( $p > 0.001$ ). No significant effects of accent origin were found for the formant excursion metrics.

## 5.5 Discussion

Assuming the validity of the vowel reduction metrics employed in other studies, the correlation results between the reduction metrics and the vowel quality metrics provide preliminary support for the hypothesis that rhythm may be reflected in vowel quality. We may conclude that more studies of this type, covering more data and more languages, are warranted.

The results here also provide some perspective on attempts to evaluate rhythm metrics. White and Mattys (2007) compare a number of metrics across different languages as well as different accents of English. These studies suggest that Varco V (and %V, a phonotactic metric) capture hypothesized rhythm differences better than PVI-V. In this connection we may notice that two of the accents with the lowest PVI scores characteristic of syllable timing, Belfast and Littlehampton, have relatively high VarcoV scores suggesting stress timing. Grouping accents according to PVI would mean placing these two samples in a category with two of the three Polish accents and one of the

Welsh ones. The VarcoV metric seems to better capture the established rhythmic characterizations of these accent groups.

The %Steady State metric is the only vowel quality metric in which significant differences were found among groups of accents. Additionally, of the vowel quality metrics, %Steady State showed the strongest correlations with the reduction metrics. These findings suggest that this metric is the most promising of the vowel quality metrics for future investigations. This indeed may be expected when we consider that the formant excursion metrics, since they do not quantify the rate of formant change, do not necessarily describe the robustness of formant movement. In other words, significant formant movement, if it is so rapid as to be almost instantaneous, may not be terribly salient. This point of course raises the question of quantification – at what point as an acoustic transition too rapid to be reliably heard? The formant excursion metrics employed here offer no insight into this issue. The %Steady metric, on the other hand, quantifies vowel purity in a more straightforward way.

## 6. General discussion

The experiment described in this paper was borne out of a view that the interaction between segmental and suprasegmental phonology represents a pressing issue for phonological theory, and seeks to build on recent attempts to describe and quantify rhythm in terms of phonological parameters such as vowel reduction and phonotactics. Data from various accents of English suggest that vowel quality, and in particular vowel purity quantified with a %Steady State measure, may be linked with phonological vowel reduction. Vowel reduction is an interesting phenomenon in phonology since it lies at the interface between segmental features and rhythm. The possibility that vowel quality is tied to this interface as well may serve as useful building block in the construction of a holistic model of phonology in which the interaction between segmental and suprasegmental phenomena becomes transparent. In particular, we shall concentrate on how rhythmic issues may tie in with the representation of vowel quality. This in turn is closely linked with cross-accent variation in English, as well as obstacles in the acquisition of English as a foreign language.

For a long time the perception of vowels was widely assumed to be characterized in terms of a “simple target model” (for a review, see Strange 1989), in which formant values from quasi steady-state portions were taken to be acoustic targets for speakers. In the presence of acoustic variability caused by coarticulation with neighbouring consonants, listeners were said to compensate for these effects and reconstruct canonical values associated with the targets. Findings in Strange et al (1983), reconstructed in a number of later studies, provided a serious challenge to the simple target model. In these experiments, listeners were found to be able to identify coarticulated vowels *MORE* accurately than vowels produced in isolation. However, these findings, which spawned a “dynamic specification” theory of vowel perception in which formant trajectories play a pivotal role in listener identification, were largely limited to studies of English. As a result, dynamic specification may be a language-specific phenomenon, predicted by the present experiment to predominate in stress-timed languages.

Dynamic specification in English vowels has significant implications with regard to the “target” articulations of traditional transcription. We may speculate that such dynamic specifications result in variability in the temporal location at which the vowel is closest to its target. For example, the /i:/-/ɪ/ contrast may be represented as a difference in target location. The lax /ɪ/ target is typically reached early in the vowel. Tense /i:/, on the other hand, frequently shows a pattern in which the extreme high front target is reached toward the end of the vowel. This difference is represented schematically in Figure 2, which represents the time windows occupied by the two vowels, /i:/ on the left and /ɪ/ on the right, with target location specified along the timeline.

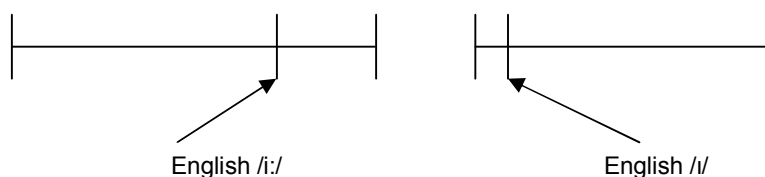


Figure 2 – Schematic representation of the temporal aspects of the tense-lax contrast in English

In accents and languages with high values for the % Steady State measure, vocalic targets must be reached early. Otherwise, a “pure” vowel would be impossible. Figure 3 projects the English contrast from Figure 2 alongside a similar representation of Polish /i/, which is typically quite pure in quality. These representations provide a provocative perspective on what might constitute a foreign accent when a speaker of a syllable-timed language with pure vowels attempts to produce a long and tense /i:/. For example, when the word *peace* is spoken with a heavy Polish accent, native English speakers often hear *piss*.

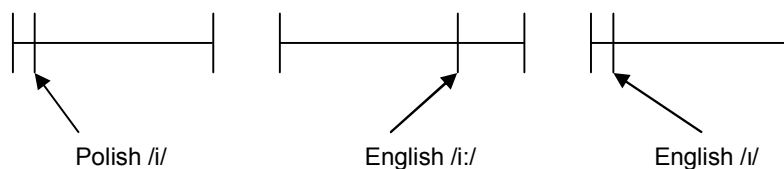


Figure 3 – Differences in dynamic specification between Polish and English /i:/

From a perceptual perspective it is likely that problems distinguishing *peace* from *piss* have less to do with tongue position or duration of the vowels than it does differences in their dynamic properties. If Polish speakers base vowel identification decisions on formant patterns early in the vowel, as suggested by findings in Schwartz (2007), mishearing *peace* is to be expected, since at the key perceptual moment for Polish listeners, the vowel has not yet reached its /i/-like target. The representations offered in the two figures capture the dynamic properties of English /i:/ that are often elusive to Polish learners. They are intuitive and easy to interpret, offering learners an additional tool for vowel acquisition that traditional vowel charts and segmental transcription are unable to provide

An additional phonological link in the chain that underlies the hypothesis explored in this study is consonant lenition. The basic premise is that when listeners may reconstruct consonants on the basis of dynamically specified vowels, weakening has lesser perceptual consequences than it would when a lesser portion of pure vowel is available for consonant identification. Widespread lenition in a language implies significant variability in consonant strength. Acoustic properties associated with strength have been found to affect the temporal location of p-centres, perceptual “moments of occurrence” of syllables (Scott 1993). As a result greater variability in strength should result in variability in p-centre location that is expected in stress-timed languages.

Experiments on p-centre location have not, to my knowledge, investigated the possibility of cross-language differences. Many aspects of speech perception are language specific and dependent on the phonology of the given language (e.g. Boersma and Hamann 2009). Thus, it is conceivable that listeners adjust p-centre location in accordance with the language's phonological specification for rhythm. We might therefore predict that in languages with low VarcoV and high % Steady State values, the effects of onset strength on p-centre location are less significant. Rhythm is largely a perceptual phenomenon, and perception is subject to language-specific phonological constraints. As a result, a promising approach to rhythm would be to consider it in segmental rather than suprasegmental terms. This strategy offers a number of benefits for both theoretical description as well as language learning.

## References

- Abercrombie, D. 1967. *Elements of general phonetics*, Chicago: Aldine.
- Bertinetto, P. 1989. Reflections on the dichotomy 'stress' vs. 'syllable-timing'. *Revue de Phonetique Appliquee*, 91-93, 99-130.
- Boersma, P. and S. Hamann. 2009. Loanword adaptation as first-language phonological perception. In Calabrese, A. and W. L. Wetzels (eds.), *Loanword phonology*. Amsterdam: John Benjamins. 11-58.
- Dellwo, V. 2006. Rhythm and speech rate: A variation coefficient for deltaC. *Language and Language Processing: Proceedings of the 38<sup>th</sup> Linguistic Colloquium* ed. by Pawel Karnowski & Imre Szigeti. 231-241. Frankfurt: Peter Lang.
- Donegan, P. and D. Stampe. 1983. Rhythm and the holistic organization of language structure. In *Papers from the Parasession on the Interplay of Phonology, Morphology, and Syntax*. Chicago: CLS 19, 337-353.
- Hume, E., K. Johnson, M. Seo, G. Tserdanelis and S. Winters. 1999. A cross-linguistic study of stop place perception. In *Proceedings of the 14th International Congress of Phonetic Sciences*, 2069-2072.
- Jenkins, J. 2000. *The phonology of English as an international language*. Oxford: OUP.
- Low, L. E.; E. Grabe and F. Nolan. 2000. Quantitative Characterizations of Speech Rhythm: Syllable-Timing in Singapore English, *Language and Speech*, 43(4), pp.377- 401.
- Pike, K. L. 1945. *The intonation of American English*. Ann Arbor, MI: University of Michigan Press.

- Ramus, F., M. Nespore and J. Mehler. 1999. Correlates of linguistic rhythm in the speech signal. *Cognition* 73, 265–292.
- Schwartz, G. 2007. Vowel quality and its holistic implications for phonology. Ms. Adam Mickiewicz University.
- Scott, S. 1993. P-centres in speech – an acoustic analysis. Ph.D. Dissertation, University College London.
- Strange, W. 1989. Evolving theories of vowel perception. *J. Acoust. Soc. Am.* 85. 2081–2087.
- Strange, W., J. Jenkins and T. Johnson. 1983. Dynamic specification of coarticulated vowels. *J. Acoust. Soc. Am.* 74, 695–705.
- Wagner, P. 2007. Visualising levels of rhythmic organization. *Proceedings of ICPhS XVI. Saarbrücken.*
- White, L. and S. L. Mattys. 2007. Rhythmic typology and variation in first and second languages. In Prieto, P., J. Mascaró and M.-J. Solé (eds.), *Segmental and Prosodic issues in Romance Phonology*. Current Issues in Linguistic Theory series (pp. 237–257). Amsterdam: John Benjamins.
- Wright, R., S. Frisch and D. Pisoni. 1997. Speech Perception. Research on Spoken Language Processing, Progress Report No. 21. Indiana University.