Poznań Studies in Contemporary Linguistics 46(4), 2010, pp. 429–456

doi:10.2478/v10010-010-0022-6

TOWARDS A GRADUAL SCALE OF VOWEL REDUCTION: A PILOT STUDY

MAŁGORZATA KUL
Adam Mickiewicz University, Poznań
kgosia@ifa.amu.edu.pl

ABSTRACT

The study reports the results of an acoustic analysis of vowel reduction of the /iː/ vowel, considering all three traditionally explored aspects of vowel reduction, i.e. duration, F1 and F2 in read speech produced by 12 native speakers of English. Starting from the observation that the standard literature considers only duration as a proxy for overall reduction, the aim of the study is to verify whether duration, F1 and F2 exhibit reduction (construed as shortening of duration and centralization of formants, respectively) to the same degree. The *r* test reveals the lack of a robust linear correlation between duration, F1 and F2, the highest value being 0.51 (the correlation between duration and F1) and 0.24 (the correlation between duration and F2), neither of which is a strong correlation. In light of the results, the study seeks to establish a gradual scale of vowel reduction, combining the spatial and the temporal aspects by means of averaging the distances between the least and the most reduced tokens across duration, F1/F2 on an equal basis. The resulting degree is expressed on a scale of reduction, ranging from 0 (no reduction whatsoever) to 100 per cent (reduction to schwa).

KEYWORDS: Vowel reduction; quantifying reduction; normalized ratio; combined parameters approach; gradual scale.

1. Introduction

The mechanisms underlying phonetic reduction have long been noted, widely investigated and extensively documented across languages. With regard to vowel reduction, numerous studies offer an account of the phenomenon, ranging from centralization (Koopmans-van Beinum 1980) through undershoot (Lindblom 1963) to information loss (Harris 2005). More recently, the factors triggering reduction have held a position of particular prominence in the literature. The factors (both linguistic and nonlinguistic) identified so far include frequency (Jurafsky et al. 1998; Bybee and Scheibman 1999; Bell et al. 2003; Hollmann and Siewierska 2007; Zhao and Jurafsky 2009; Bell et al. 2009), lexical vs. function word status (Kelly and Local 1989; Lavoie 2002; Corbin 2003; Local 2003), old vs. new information (Howes 1967; Fowler and Housum 1987;

Greenberg and Fossler-Lussier 2000; Mitterer and Ernestus 2006; Bradlow and Baker 2009), rate (Shockey 2003), stress assignment (Port 1981), spontaneous vs. read speech (Barry and Andreeva 2001), position in the syllable (Ohala and Kawasaki 1984), duration (Max and Caruso 1997), intraspeaker variability (Ellis and Hardcastle 2002; Hollmann and Siewierska 2006), interspeaker variability (Byrd and Tan 1996; Piroth and Janker 2004), audience design (Beckford Wassink et al. 2007), manner of articulation (Byrd and Tan 1996; Saltzmann 2003; Recasens 2004), environment in terms of following and preceding sounds (Duez 2001; Mitterer and Ernestus 2006), sentence stress (Lennes et al. 2001), rhythmic group (Barry and Andreeva 2001), the (in)alienability effect (Hollmann and Siewierska 2007), sex (Byrd and Tan 1996; Bell et al. 2003; Tagliamonte and Roeder 2009) familiarity with speaker (Newman and Evers 2007), background noise (Rosenhouse and Rabin 2003; Zhao and Jurafsky 2009), coarticulation (Ellis and Hardcastle 1999), type of lenition process (Clark and Trousdale 2009), idiomaticity (Sanford 2008), consonant-cluster type, e.g. cross-boundary clusters (Barry and Andreeva 2001), predictability (Bell et al. 2003), and planning problems/ dysfluencies (Jurafsky et al.1998).

Usually, a single factor causing reduction is investigated at a time; yet a growing body of work is exploring the possibilities of an interaction between the factors: “Although there is a sizable body of research on the effects of probability and speech style on articulation level, few studies have examined how such factors interact with each other” (Bradlow and Baker 2009: 395). Instead of holding e.g. frequency as the sole factor responsible for reduction, these studies combine for example frequency and predictability (Jurafsky et al. 1998; Bell et al. 2003; Bell et al. 2009) or frequency and noise (Zhao and Jurafsky 2009), seeking to establish the role of individual factors in relation to each other. In the face of these advances, a sufficiently fine-grained and comprehensive method of measuring reduction appears to be highly relevant to the success of these attempts. A brief overview of the recent literature in the field reveals two major trends with regard to measuring reduction: either duration of a word (Jurafsky et al. 1998; Bell et al. 2003; Sanford 2008; Bell et al 2009; Baker and Bradlow 2009) or of a vowel (Aylett and Turk 2004) is taken to quantify reduction. Work by Aylett and Turk (2004, 2006) is the only notable exception to the tendency since it considers both durational and spatial effects of centralization. These two measures, however, are used for different studies and different vowel database. Tagliamonte and Roeder (2009) claim to also consider intensity, but actually take this no further. The traditional narrow ways to measure reduction in vowels raise a number of issues.

Firstly, a sizeable body of the phonetic literature successfully evaluates the effects of reduction in vowels by means of investigating spectral differences between citation and reduced forms (Picheny et al. 1986; Lindblom 1990; Moon and Lindblom 1994; Bradlow et al. 1996; Aylett and Turk 2006; Mooshammer and Geng 2008), differences due to prosody (Summers 1987; van Bergem 1983) or attributable to word frequency (Bybee 2000; Wright 2003; Munson and Solomon 2004). With regard to the above studies, it appears that the studies in casual speech ignore formants and, by the same token, phonetic tradition where formants, not duration, reflect vowel quality best. The only study, to the best of my knowledge, assessing the temporal vs. spectral aspect is by Aylett and Turk (1998). This study, however, concerns perception of the vowel's goodness (defined as perceptual clarity) rather than production-connected, actual aspects of reduction. Yet, in a more speaker-oriented study, Aylett and Turk (2006) do find a significant relationship between language redundancy factors (i.e. probability) and the first two formants of the vowel: "spectral characteristics (e.g. formant frequencies in the central portion of a vowel, among other things) can also be considered as measures of acoustic redundancy, since these are known to relate to vowel distinctiveness and thus probability of recognition" (Aylett and Turk 2006: 3048). Thus, the decision to employ only duration of a vowel in quantifying reduction seems arbitrary and devoid of empirical evidence.

Secondly, most studies (Jurafsky et al. 1998; Bell et al. 2003; Sanford 2008; Bell et al 2009; Baker and Bradlow 2009) consider reduction in terms of duration of whole words (or, less frequently, syllables), although vowels exhibit reduction in a more reliable way than words. In the phonetic tradition, duration of the vowel and not of a syllable or a word is employed to signify the changes across speaker's populations (e.g. Hillenbrand et al. 1995; de Jong et al. 2007; Mooshammer and Geng 2008). This tendency seems to be especially problematic in the situation where a high degree of precision is required to uncover the relative significance of the factors behind reduction.

Thirdly, the few studies that do consider vowel duration in pursuing reduction determinants (e.g. Lavoie 2002), appear to assume that temporal changes automatically entail spatial ones. In fact, it is generally assumed that for vowels, formant frequency does neatly correlate with duration. Even though this relationship is undisputed in the literature, the exact nature of the relationship between vowel duration and frequency has never been a matter of a separate debate (Moon and Lindblom's 1994 work being a rare exception), the implicit assumption being that the relationship is of a linear character. This assumption derives from the theory of vowel centralization. The suggestion that a stress-induced decrease in duration involves a tendency towards centralization can be traced back to Joos (1948), Tiffany (1950), Shearme and Holmes (1962), Delattre (1965/1981) Fant et al. (1974) or Koopmans-van Beinum (1980). In other words, the centralization theory posits that a decrease in duration of a vowel entails the changes (i.e. an increase or a decrease, depending on the quality of the vowel in question) in the values of F1 and F2, which approximate the vowel to schwa, and operates on the principle that the relationship between duration and formant frequencies is linear. For instance, if the vowel /ɪ/ is being centralized, it means that with a decrease in duration, its F1 increases (from 370 Hz to roughly 500 Hz of schwa, cf. Cruttenden 2008: 99), whereas its F2 decreases (from 1700 Hz to approximately 1400 Hz of schwa, cf. Cruttenden 2008: 99).

In an attempt to empirically test the linearity hypothesis, this study seeks to explore the nature of the relationship between duration, F1 and F2. The first two formant frequencies of a vowel were selected due to their well-established status as indicators of reduction (see references above) or vowel quality: "Formants have long been held to be well-suited parameters for describing vowel quality, mainly because the F1/F2 plane correlates with traditional articulatory–auditory trapezoidal representations of vowels" (Ferragne and Pellegrino 2010: 4). Therefore, the paper sets out to verify whether the current studies on the factors triggering reduction are justified in considering only the temporal aspect of vowel reduction. In this respect, the study fleshes out a suggestion made by Baker and Bradlow (2009: 408), who also see the focus on just one aspect of reduction as a potential problem: "This work has opened up a new avenue of research examining the interaction between different types of probability factors and speech styles. One possibility is to look at how lexical probability and speech style affect articulatory features other than word durations such as vowel space and amplitude".

2. Method

2.1. Predictions

Given that the current literature on reduction considers only the duration of a vowel (or even less reliably, of a syllable or of a word), the paper sets out to test the following predictions concerning duration, F1 and F2 (as the first two formants of the vowel are usually taken into account):

Prediction 1: Duration, F1 and F2 are linearly correlated.

Prediction 1a: Duration is linearly correlated with F1.

Prediction 1b: Duration is linearly correlated with F2.

2.2. Speech material

In order to address the paper's research question, an acoustic study of reduction in vowels was designed. The study undertakes a spectral analysis of the vowel /iː/. The vowel /iː/, among all tense vowels, is reported to be the most stable and to exhibit the most consistent spectral and durational changes: "Groups 1 (high language redundancy and 3 (low language redundancy) differed in a direction consistent with the hypothesis that language redundancy correlates with centralization, with /i/ showing a particularly good match to prediction" (Aylett and Turk 2006: 3055). A range of realizations of /iː/ were elicited, varying widely in their degree of reduction due to repetitions of the carrier word (*key*). Since the study focuses on vowels, /k/ was not considered (but see Section 5 for some suggestions on how the method can be extended to consonants). The database consisted of samples of the vowel /iː/ obtained from elicited speech recordings of native speakers of English.

2.3. Subjects

Twelve subjects, ranging in age from 25 to 61, represented various dialects of the English language, i.e. American, Irish, British and Canadian. None of the subjects were familiar with the aims of the study.

2.4. Text

All subjects read the same text. Although the vowel /iː/ was embedded in a stressed syllable, the effect of vowel reduction was induced by including subsequent tokens of the same word in a text (Howes 1967; Greenberg and Fossler-Lussier 2000; Anderson and Howarth 2002; Mitterer and Ernestus 2006; Bradlow and Baker 2009). In order to control for the context factor, every token of the vowel /iː/ was derived from the phrase *the key player*, embedded in carrier sentences. In order to fully isolate the reduction effects from coarticulation, /iː/ in the word *team* was excluded from the analysis (see Cruttenden 2008: 95 for details about the duration of /iː/ followed by a nasal as against other segments). The text was as follows:

> *The key player was absent in that match. I was sorry for his team because the key player has always scored most goals. Unfortunately, the key player's team lost the championship. It proved that the key player was indispensable.*

2.5. Recordings

In the text, there were four tokens of the word *key* in the context of the phrase *the key player*, while the recordings of each of the twelve subjects were conducted seven times. This number results from calculating the modified coefficient of variation. The modified coefficient of variation (the coefficient of variation for the mean value), established on the basis of the sample size, for an analyzed parameter (here the vowel duration parameter), is the relation of standard deviation of the mean value to the mean value, given in per cent (Oktaba 1966) and whose results, calculated for duration if /iː/, are displayed in Figure 1, below.

Figures 1 and 2 illustrate the modified coefficient of variation for two different speakers (Speaker 0 did not participate in the experiment and served to independently verify the sample size). It indicates that the desired sample size is 7 readings per speaker, as beyond the seventh reading, the results did not change anymore. Thus, the number of tokens of /iː/ in the database totalled 336 (twelve speakers, four tokens, seven readings).

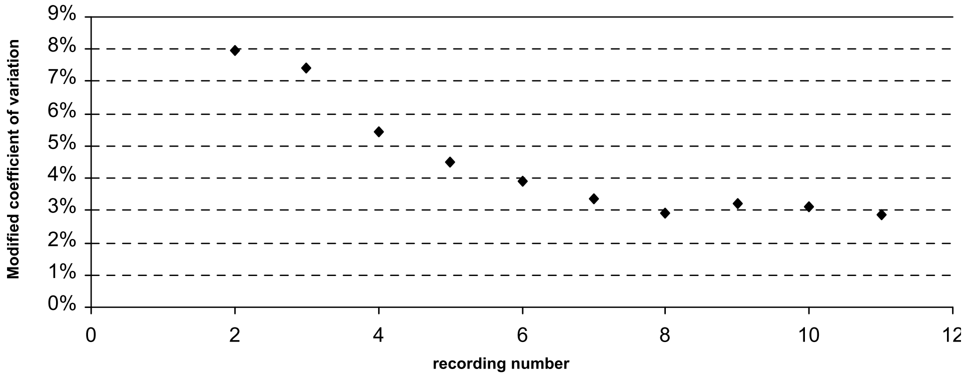

Figure 1. The sample size (for Speaker 1).

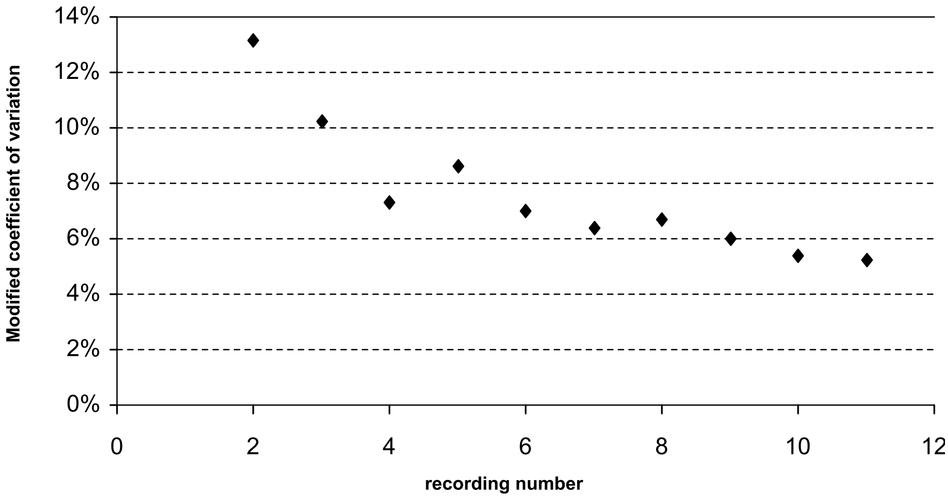

Figure 2: The sample size (for Speaker 0).

2.6. Measurements

Using the Praat software, spectrograms of all tokens of the vowel /iː/ were produced, segmented and annotated manually. Following established practice, conservative criteria were employed to determine the onset and the offset of the vowel. They include the presence of vowel formants (following a noise burst attributable to the preceding stop and read from the spectrogram), and regularity of the waveform (as the C to V transition assumes a regular shape, derived from the waveform). Voicing was excluded from the analysis as the vertical striation (as well as blue lines signifying pulse in Praat) may well result from the lowest degree of background noise. Next, the midpoints of the vowels were judged by visual inspection. At midpoint, measurements of F1 and F2 were taken manually (i.e. without resorting to a Praat script, which is normally only used for larger samples). Besides, it is a more reliable way than any automatic procedure as "to date, no automatic tool can reliably extract formants" (Ferragne and Pellegrino 2010: 4). Measurements of duration were taken from the onset to the offset of the vowel.

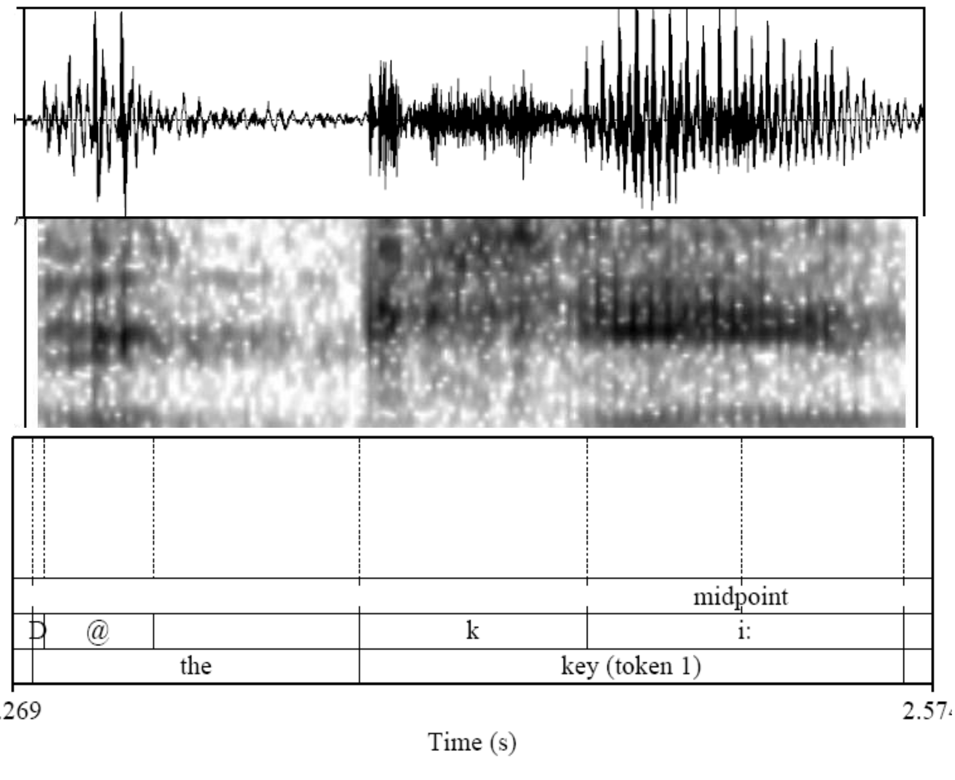

Figure 3: An example of a Praat annotation (Speaker 4).

Figure 3 captures a sample annotation, depicting the way the recorded speech was segmented. Then, the obtained values were averaged across seven repetitions of the text for every speaker with regard to the individual parameters of reduction, i.e. duration, F1 and F2. Finally, the Pearson's significance correlation test was performed on the three parameters of vowel reduction.

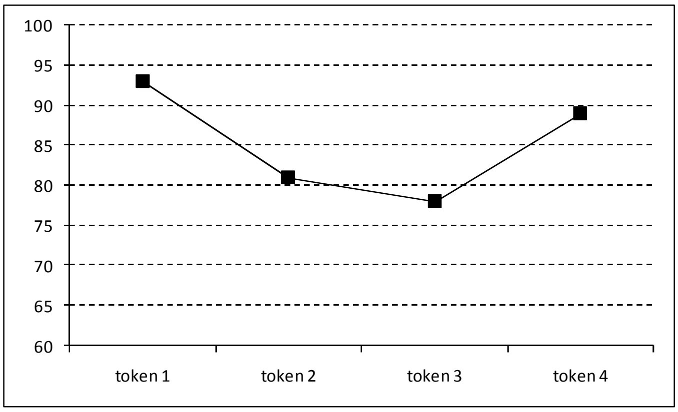

Figure 4. An example of reduction for duration (msec) averaged over seven repetitions (Speaker 10).

3. Results

3.1. Results for the overall correlation between duration, F1 and F2 (Prediction 1)

It was hypothesized that the new information contained in the first, introductory token (token 1) of the word *key* would be reduced in its subsequent tokens, numbered 2, 3 and 4, carrying old/given information (Fowler and Housum 1987). This hypothesis derived from the Hyper- & Hypospeech theory (Lindblom 1990) according to which the speaker adjusts his articulation according to the needs of the listener. It implies that token 1, as the introductory one, would resist reduction, whereas tokens 2, 3 and 4 would undergo reduction to various degrees. At issue is whether duration, F1 and F2 equally reflect reduction exhibited by tokens 2, 3 and 4 relative to token 1 (the question formulated in Prediction 1).

According to the H&H-based hypothesis, from token 1–4, a progressive approximation in terms of F1 and F2 towards the "typical" values of schwa was expected. "Typical" is defined here following Cruttenden (2008: 99), i.e. for male speakers (the gender of all subjects but one), 513 Hz and 1377 Hz for F1 and F2, respectively. Over the seven readings, the average value of the speakers' first token was indeed seen to be the least centralized, compared to the values of the "typical" schwa.

The differences observed in duration, F1 and F2 across the four tokens are statistically significant (p=0.025). Throughout the study, one-tailed tests were used since a directional hypothesis was tested (Section 2.1). Note, though, that the significance of this effect does not entail that duration, F1, and F2 change in exactly the way that the hypotheses predicts. It merely suggests that there are differences *of some kind* across token 1–4 which are unlikely to be the result of chance. The specific Predictions 1, 1a and 1b

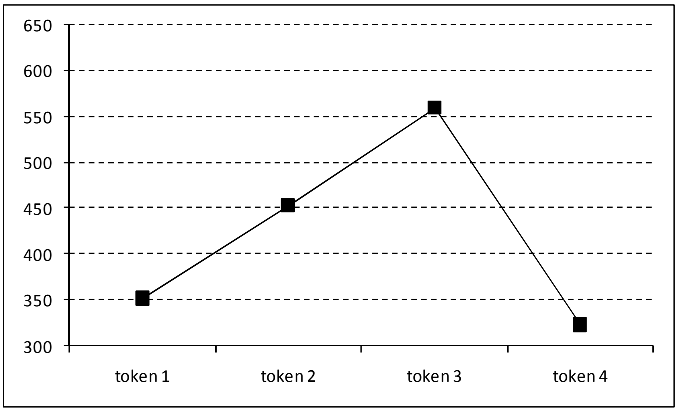

Figure 5. An example of reduction for F1 (Hz) (Speaker 7).

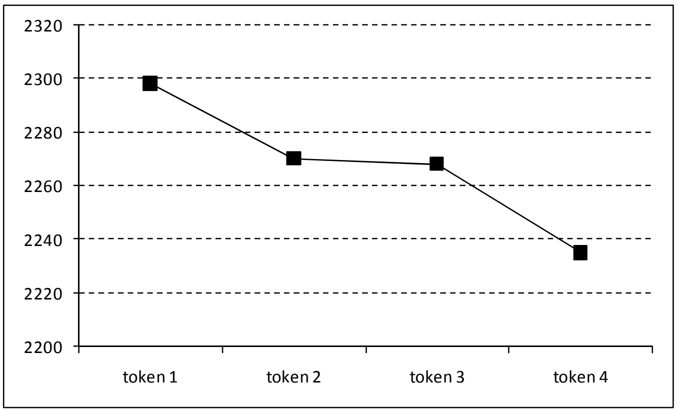

Figure 6. An example of reduction for F2 (Hz) (Speaker 12).

are tested below. Thus, the results for Prediction 1 are visualized in Figures 4, 5 and 6, respectively.

It is apparent from Figure 4 that the speaker fails to exhibit a consistent propensity towards reduction since token 4 is hyperarticulated. This can be explained by the structure of the text where token 4 is embedded in the last sentence of the text, in which the speakers' attention might possibly rebound. To test this hypothesis, I carried out an extra measurement on one speaker (Speaker 0, the speaker who verified additional measurements in the study but did not participate in it) in which I added another sentence, following the last one, to the text. This eliminated the hyperarticulation effect from token 4, providing evidence for an attention-related explanation; see Figure 7.

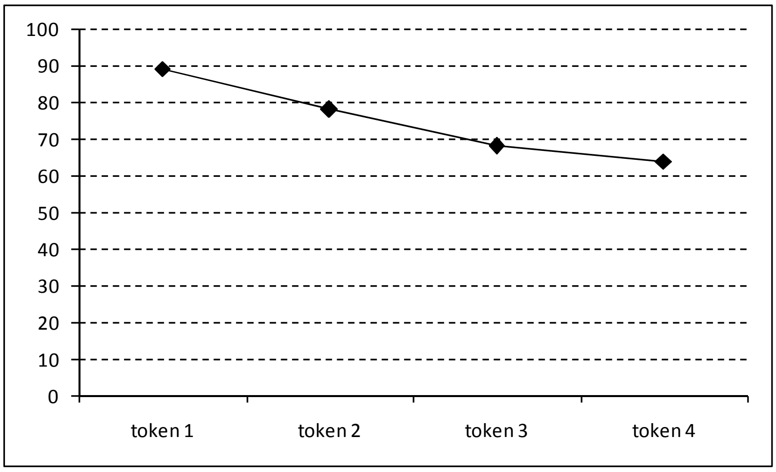

Figure 7. Duration of the /iː/ vowel in subsequent tokens of *key* (msec), with token 4 followed by an extra sentence (Speaker 0).

A more important finding perhaps is that in studying the patterning of reduction across parameters, a great deal of interspeaker variability was observed. Only five speakers out of twelve demonstrated consistent changes across all the parameters of reduction (Speakers 4, 5, 7, 8, 9). In other words, not all the speakers reduced all the parameters simultaneously. The degree of interspeaker variability exhibited in reduction across parameters is captured by Table 1 below. The signs denote lack or presence of consistent changes across tokens (i.e. a decrease in duration, an increase in F1 and a decrease in F2 values).

Table 1. Interspeaker variability in reduction across parameters.

Speaker no.	Duration	F1	F2
1	+	+	
2	+		
3	+	+	
4	+	+	+
5	+	+	+
6	+		+
7	+	+	+
8	+	+	+
9	+	+	+
10	+	+	
11	+	+	
12	+		+

This is the first piece of evidence to undermine the traditional assumption of duration, F1, and F2 all being neatly correlated in reduction processes. It emerged that 100 per cent of speakers shortened duration in tokens 1–4, 75 per cent of speakers increased F1, whereas 58 per cent of speakers decreased F2. These findings point to consistent changes in duration, followed by fairly robust changes in F2, meanwhile changes in F1 were not that considerable.

Turning to the results of the study for Prediction 1, they are summarized by Table 2 below.

Table 2. Average values of reduction of the tokens across parameters (for all speakers).

	Token 1	Token 2	Token 3	Token 4
Duration (msec)	87.15	74.77	68.30	74.33
F1 (Hz)	415.12	427.63	413.81	373.72
F2 (Hz)	2284.45	2280.12	2292.90	2273.43

These values can be compared across parameters as represented by Figure 8.

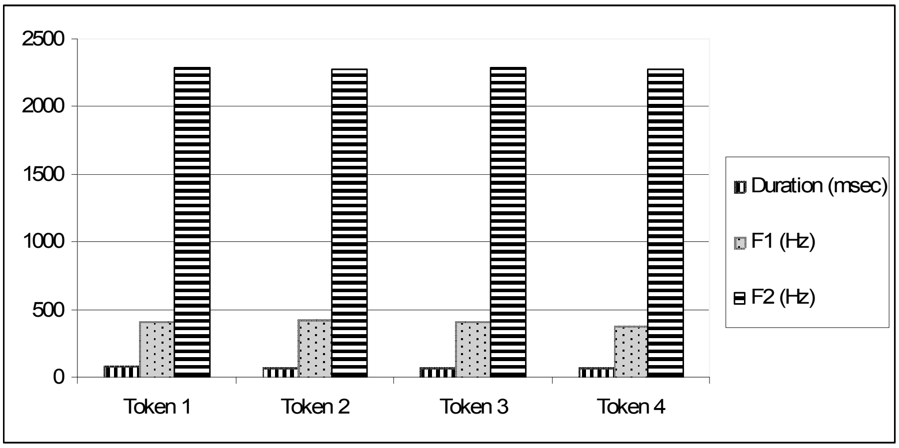

Figure 8. Cross-parameter comparison of reduction
(average values of reduction for all speakers).

In Table 2, the columns give the values obtained from the study for the parameters of vowel reduction, averaged across the speakers. As was expected, duration decreases across subsequent tokens, relative to token 1, although it bounces back up again a bit in token 4. The values of F1 should increase (representing centralization), which was observed only for token 2. The values of F2 ought to decrease, but this was not noted for the third token.

One should note that the values measured for both F1 and F2 are rather high (see Table 1) in comparison with the newest formant values such as those given by Ferragne and Pellegrino (2010). In their study, F1 ranges from 248 to 335 Hz (across thirteen accents of the British Isles), whereas F2 ranges from 2164 to 2346 Hz (Ferragne and Pellegrino 2010). This difference can be ascribed to a high degree of influence of preceding /k/. As a velar consonant, it has caused a significant increase of F2 values. However, by controlling for the context (i.e. embedding the vowel /iː/ in the same context) we can still observe the changes in formant frequencies independent of the contextual effects. Interestingly, the changes in F2 proved to be slightly less robust than the changes in F1, which was also documented by Aylett and Turk (2006) (Table V, p. 3056) and de Jong et al. (2007): "For both /uː/, the vowel in WHO'D, and /ʊ/, the vowel in HOOD, the frequency of F2 has increased considerably, indicative of a more fronted pronunciation of those vowels, while the F1 has remained relatively unaffected" (de Jong et al. 2007: 1815).

This is perhaps unsurprising because phoneticians traditionally assume that F1 is perceptually more salient than F2 due to the fact that, at least as far as the cardinal vowels are concerned, the former involves five categories (high, mid-high, mid, mid-low and low) whereas the latter features only three (back, central and front). It has been established that, in comparison with a weighted average of F2, F3 and F4, perceptual salience of F2 is smaller (Carlson et al. 1975).

As is apparent from Table 2, while duration undergoes reduction to the greatest degree, formant frequencies do not manifest an equal degree of reduction. In fact, token 4 was not reduced (centralized) in relation to its F1, while token 3 both F1 and F2 were not reduced, which stands in stark contrast to the pattern displayed by duration. This finding that duration, F1 and F2 fail to exhibit consistent changes in reduction across parameters, can be also presented graphically.

Table 3. Differences across parameters in showing reduction in subsequent tokens (2–4) relative to token 1 (for all speakers).

	Token 1	Token 2	Token 3	Token 4
Duration (%)	0	15	22	15
F1 (%)	0	3	1	10
F2 (%)	0	1	0	1

Figure 9 provides an overview of the trends, depicting the differences across the parameters of reduction. The figure clearly supports the observation that reduction exhibited by duration trends in a different direction than reduction manifested in F1 and F2.

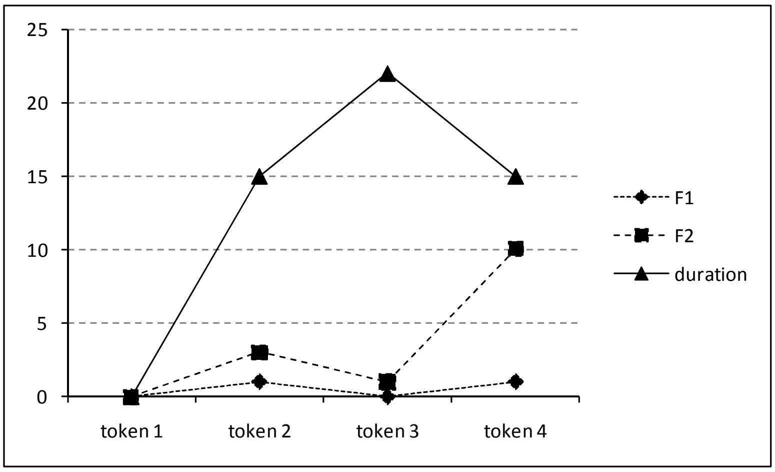

Figure 9. Degree of reduction of tokens 2–4 relative to token 1 across parameters (for all speakers).

Thus, in the light of the evidence that the spatial and the temporal aspects of reduction do not neatly pattern together, Prediction 1 cannot be vindicated.

3.2. Results for individual correlations (Predictions 1a and 1b)

(1) Prediction 1a: duration is correlated with F1. This prediction assumes that any decrease in duration will also involve an increase in the values of F1 across tokens (as F1 is inversely related to the height of the tongue).

(2) Prediction 1b: Duration is correlated with F2. This prediction stipulates that any decrease in duration will also involve a decrease in the values of F2 across tokens (as F2 is proportionally related to the position of the tongue).

Predictions 1a and 1b were not verified on all twelve speakers who participated in the experiment reported in this paper due to a considerable degree of interspeaker variability in employing strategies of reduction (see Table 1). Instead, data on vowel reduction obtained from only five speakers were used (Speakers 4, 5, 7, 8, 9). Only these speakers changed all the parameters simultaneously (which is not to say that they were neatly correlated; see further Section 3.2). The remaining seven changed only one or two of the three parameters at the same time, contra the traditional assumption. These speakers were therefore excluded from the correlation analysis.

Prediction 1a–b stipulated the existence of a linear correlation between parameters of reduction in vowels. To this end, the presence and strength of a possible correlation was investigated using a simple Pearson correlation coefficient. Table 4 below presents the results, along with the *p* values:

Table 4. Pearson's correlation coefficient averaged across tokens (for Speakers 4, 5, 7, 8 and 9).

	Duration–F1		Duration–F2	
	r value	p value (one-tailed)	*r* value	p value (one-tailed)
Token 1	0.35	0.001	0.37	0.016
Token 2	0.24	0.08	0.25	0.08
Token 3	−0.12	0.24	−0.09	0.032
Token 4	0.52	0.008	−0.12	0.26
Average	0.25		0.10	

The numbers presented in Table 4 do not support the linearity hypothesis due to low or negative *r* values, signifying a lack of correlation between the respective parameters of reduction. The average scores of duration and F1 and duration and F2 are visually displayed in scatter plots in Figures 10–11.

As it is apparent from Figures 10 and 11, the correlation between duration and F1 is not linear, as the values fail to achieve a linear arrangement. Neither is the correlation between duration and F2. The correlations between duration and F1/F2 are not significant; the results indicate a lack of or an extremely feeble correlation at best.

4. Discussion

The study of the currently employed ways of quantifying reduction in vowels has provided clear answers to the research questions. Firstly, the finding that duration, F1 and F2 do not reflect the same changes induced by reduction means that we must reject Prediction 1. The observed lack of uniformity in the parameters of vowel reduction across speakers can be explained as either individual strategies arising from interspeaker variability (Ferragne and Pellegrino 2010) or may perhaps be seen as a consequence of read speech.

Secondly, the lack of a significant correlation between duration on the one hand, and F1 or F2 on the other suggests the lack of a linear relationship between these pairs of variables. This result seems to run counter to one of the corollaries of the centralization theory stipulating that F1 increases along with a decrease in duration, and that F2 decreases along with duration (at least for the vowels approaching the quality of schwa). The lack of a significant correlation between duration, F1 and F2, however, is difficult to interpret, as with any null result. One possibility is that the relationship is other than linear,

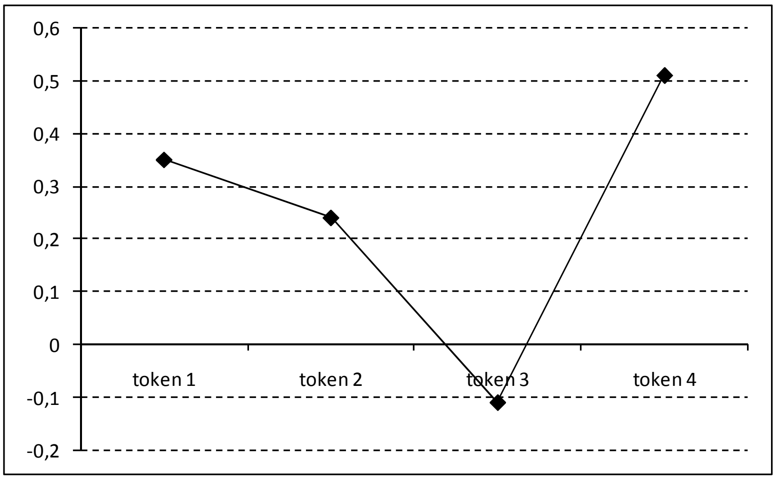

Figure 10. Scatterplot of correlation between duration and F1 (for Speakers 4, 5, 7, 8, 9).

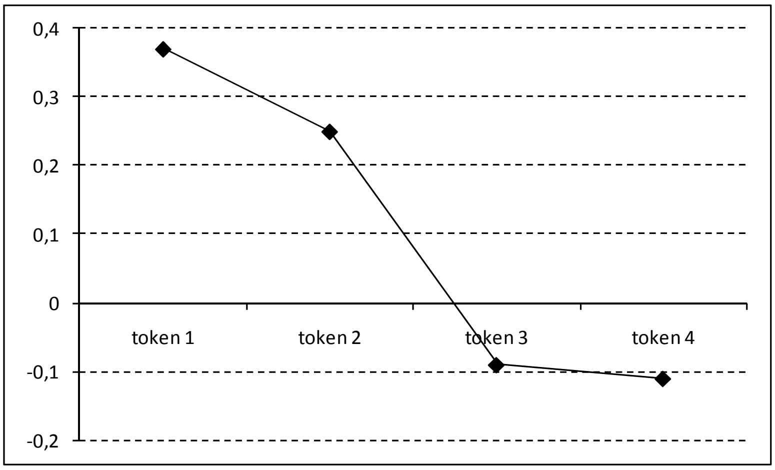

Figure 11. Scatterplot of correlation between duration and F2 (for Speakers 4, 5, 7, 8, 9).

perhaps partly because a non-linear correlation has not been explored yet. An alternative account for the findings of the study presented in an attempt to verify Prediction 1a and 1b might posit the existence of such correlations in fully spontaneous speech. In this study, the speakers read the text whereas several studies argue that reduction patterns differently in read vs. fully spontaneous speech, e.g. Engstrand and Krull (2002) (however, for the counterargument that reduction processes, unlike their frequency, do not differ across speech styles, see Shockey 1974; Barry and Andreeva 2001; Bondarko et al. 2003; Silva et al. 2003 and Jaworski 2008). Given that some studies do report findings signifying that read speech varies widely from spontaneous speech, it would still seem worth testing the linearity hypothesis in fully spontaneous speech.

In discussing the findings of the study, a major observation, potentially interesting for studies in casual speech, emerges: that by considering only duration, or only formant frequencies, the overall picture of reduction may not be seriously incomplete.This would seem to necessitate certain modifications in the current ways of measuring reduction in vowels.

Summing up, the study indicates that reduction measured as the duration of a vowel cannot be taken as a proxy for reduction measured on the basis of formant frequencies of a vowel. Nor, according to the results of the study, do there exist linear relationships between the three individual parameters, duration and F1/F2. Putting these findings together, it appears reasonable to suggest that a need to develop an alternative way of measuring reduction arises. The alternative method should be non-reductionist, i.e. based not only on duration but also on centralization (changes in F1 and F2).

5. An alternative way of quantifying reduction

The present study is unique in offering statistical analysis of acoustic measurements to problematise the established practice of studying reduction based on just one of its parameters (usually duration). However, in more general terms, issues have been raised before concerning established approaches to reduction. For example, Amir and Amir (2007: 849) suggest that

> [m]any studies have been performed to determine the influence of various factors on the degree of vowel reduction. These studies have confronted two typical problems: (a) finding suitable metrics for quantifying the degree of reduction, and (b) obtaining reliable data to represent the "typical production" of the target vowels.

In the literature, reduction is typically expressed as an ordinal variable. The number of the categories (i.e. degrees of reduction) ranges from three (basic, other full or reduced – Lavoie 2002; full, other, reduced – Jurafsky et al. 1998); four (maɪ, mi, ma, mə in Hollmann and Siewierska's 2007 study of first-person singular possessives in the Lancashire dialect) to five (a laryngeal or glottalized form /ʔ/, a voiceless alveolar stop /t/, a voiced alveolar stop /d/, an approximant, and a vocalic form in Tagliamonte and Roeder's 2009 study of definite article reduction in York). A list of phonetic variants, denoted by these categories, can be exemplified as in Figure 12.

	Basic	Other full	Reduced
a	[eɪ]	[ʌ], [ɪ]	[ə], [ɨ]
the	[ði], [i], [di]	[ðʌ], [ðɪ],[ʌ]	[ðə], [ðɨ], [ə]
in	[ɪn], [ɪ], [ɪɾ̃]	[ɛn], [ʌn], [æn]	[ɨn], [n̩], [n]
of	[ʌv], [ʌ], [ʌvv]	[ɪ], [i], [ɑ]	[ə], [əv], [əf]
to	[tu], [tʉ], [ɾu]	[tʊ], [tɪ], [tʌ]	[tə], [tɨ], [ə]
and	[æn], [ænd], [æɾ̃]	[ɛn], [ɪn], [ʌn]	[ɨn], [n̩], [ən]
that	[ðæ], [ðæt], [æ]	[ðɛ], [ðɛt], [ðɛɾ]	[ðɨt], [ðɨ], [ðɨɾ]
I	[aɪ]	[ɑ], [ʌ], [æ]	[ə]
it	[ɪ], [ɪt], [ɪɾ]	[ʊt], [ʊ], [ʌ]	[ɨ], [ə], [ət]
you	[yu], [u], [yʉ]	[yɪ], [ɪ], [i]	[yɨ], [y], [ɨ]

Figure 12. An illustration of Jurafsky et al.'s (1998) ordinal variables.

These ordinal categories, useful as they are, are not devoid of certain limitations. One of them is the fact that they are established on the basis of phonetic transcription which, in turn, are based on crude, auditory and subjective impressions (even more so if different people perform the transcription). Corbin (2003), Lavoie (2002) and Tagliamonte and Roeder (2009) are rare exceptions to the general trend: these scholars perform an acoustical analysis, securing a more objective measure of reduction. The need to seek a way to quantify reduction beyond simple phonetic transcription has also been implicitly recognized in Hollmann and Siewierska's study of reduction of possessive pronouns in Lancashire dialect:

> the status of (maɪ) is also uncertain. Given that schwa is the vowel that requires least articulatory effort, it may be a reduced form of any or all of the three other variants. For the purpose of this article we will remain agnostic with regard to these aspects of the status of these forms – we will analyse (maɪ) as the fullest variant and the other forms as reduced, with (mə) being more reduced than (mi) and (ma).
>
> (Hollmann and Siewierska 2007: 408)

Another issue connected with the categories follows from their nature and concerns the fact that a gradual scale would capture a whole spectrum of variation in a better way than an ordinal one. These observations imply that a more fine-grained method to quantify reduction should be developed, using acoustic analysis while viewing the reduced forms as a continuum, assuming a scale rather than being set within the bounds of (not always very clearly defined) categories and aim at eliminating individual differences between researchers.

Regarding the variables involved in the exact measure of vowel reduction, the studies in the field typically draw a distinction between the temporal (duration) and the spatial (formant frequencies) aspects of vowel production. Suffice it to illustrate the tendency with a quote from Aylett and Turk (2006: 3051):

> due to the possibility of centralization behavior differing markedly between vowels, F1/F2 results are presented separately for each vowel. In addition, syllable duration was used as another measure of acoustic redundancy and allowed for comparison with results in the literature.

This quote clearly points out that frequency and duration are used as separate measures.

It must be pointed out here that any attempt to combine the parameters of vowel reductions raises an issue pertaining to individual differences in speakers. Numerous studies have taken notice of the differences between speakers such as sex and body size. Only a few of them, however, have attempted to overcome the differences. One of such efforts is the so-called General Procrustus Analysis (abbreviated to GPA), put forward in Mooshammer and Geng (2008). This is how GPA is explained:

> It bears resemblance to a method frequently used in morphometrics/zoology in order to solve the problem of the superimposition of landmarks (see e.g. Gower 1975, Rohlf & Slice 1990). The adapted version of GPA, applied in the current study, is based on transforming (translating, rotating and scaling) the individual formant or fleshpoint spaces to a consensus object, i.e. the speaker-independent solution. These transformations are iteratively applied until the least squares fit of all objects is no longer improved (Rohlf & Slice 1990).
> (Mooshammer and Geng 2008: 123)

The disadvantage of GPA, apart from the relatively high degree of complexity, is the fact that it must be reapplied to every token separately. Other methods devised to overcome interspeaker variability include measuring Euclidean distances to the center of the vowel space (Aylett and Turk 2004) and hidden Markov's models (Aylett 2001). However, their potential is questioned:

> results for all techniques were disappointing, arguably because any modeling technique introduces articulatory and perceptual assumptions about degree of distinctiveness or clarity that obscure spectral effects that are subtle by nature and potentially vary across vowels.
> (Aylett and Turk 2006: 3050).

In contrast, this study addresses the issues discussed above by means of proposing a gradual scale of vowel reduction, to be referred to henceforth as the Combined Parameters Approach (CPA). In order to ensure comparability of the speakers despite the unique make-up of their vocal apparatus, a procedure of normalization was applied. This means that rather than adding the values obtained from measuring duration, F1 and F2 (which cannot be executed anyway due to different units of measurement), the relation of the least reduced token to the most reduced token will be averaged and compared across duration, F1 and F2, respectively. In other words, the distances between the least and most reduced token, rather than crude values, will be compared. This is similar to Amir and Amir's (2007) comparison of ratios of similarities which, however, considers only the spatial aspect of reduction. Applying CPA (for now) to vowels only, the duration, F1, and F2 of reduced forms are arranged on scales that range from the values associated with the full form (i.e. the least reduced token for each individual speaker) to schwa (where the speaker's canonical schwa is established individually for the speakers as an average of 28 tokens of schwa). Schwa is equated here with maximal reduction because the vowel in the reading experiment reported on above is tense, and tense vowels are never reduced to zero. For lax vowels, on the other hand, this is possible (Lodge 1984; Shockey 2003), and so maximal reduction would be equated there with zero (schwa absorption).

Having established, for each token, the degree of reduction on each of the three normalized scales (duration, F1, and F2), the scales can be combined. Duration is analysed as making up half of the input into the measure of overall reduction, with F1 and F2 together making up the other half. This division is essentially arbitrary; for justification see below.

Table 5 presents an example CPA calculation of reduction of /iː/ in one of the speakers in the experiment described above. Calculating the degree of reduction under CPA proceeds in a number of steps: (a) through (g). They will be briefly outlined in order to illustrate the new method at work.

In step (a), the results for Speaker 5, before and after averaging, are presented (as opposed to Table 1, where only averaged results are given).

Table 5a. An example calculation of reduction degree (for Speaker 5). Step (a).

(a) The results for Speaker 5				
Duration	tkn1	tkn2	tkn3	tkn4
rdng1	135	107	95	96
rdng2	131	106	99	103
rdng3	137	115	104	109
rdng4	142	110	102	118
rdng5	129	105	97	103
rdng6	134	103	100	109
rdng7	130	117	104	102
mean	**134**	**109**	**100.1429**	**105.7143**
F1	tkn1	tkn2	tkn3	tkn4
rdng1	370.7	486.76	482.41	465.77
rdng2	380.96	472.03	500.84	468.99
rdng3	376.71	485.99	496.07	468.88
rdng4	472.17	449.75	489.48	453.27
rdng5	462.7	441.79	478.75	456.53
rdng6	471.24	551.83	512.41	472.83
rdng7	391.99	423.2	544.61	433.07
mean	**418.0671**	**473.05**	**500.6529**	**459.9057**

F2	tkn1	tkn2	tkn3	tkn4
rdng1	2457.54	2263.87	2250.29	2247.19
rdng2	2157.34	2223.45	2233.8	2249.4
rdng3	2322.62	2217.87	2207.4	2219.13
rdng4	2211.44	2212.43	2245.55	2263.66
rdng5	2283.13	2204.3	2258.18	2266.36
rdng6	2120.03	2360.85	2250.2	2288.88
rdng7	2508.23	2172.89	2307.53	2263.53
mean	**2294.333**	**2236.523**	**2250.421**	**2256.879**

Step (b) consists in selecting the most and least centralized tokens. The former concerns all the tokens of the speaker before averaging, whereas the latter represents the values for the speaker's schwa as an average of 28 tokens of his schwa. Similarly to the vowel /iː/, schwa was also controlled for context and only tokens from the phrase *the key* were considered.

Table 5b. An example calculation of reduction degree (for Speaker 5). Step (b).

(b) Selecting the most and least centralized tokens		
	token 1	schwa
Duration	142	30.60119
F1	370.7	464.64
F2	2508.23	1940.389

Having identified the most and least centralized tokens, a scale of reduction can be established in step (c), with schwa and token 1 as the most and least reduced tokens, respectively. The scale sets the extreme points for all three parameters of reduction considered here, i.e. by comparing the longest token to the shortest, the highest to the lowest and the most to the least advanced, relative to the speaker's center of their vowel trapezium. In this way, the obtained scale denotes ratios of tokens, rather than raw values.

Table 5c. An example calculation of reduction degree (for Speaker 5). Step (c).

(c) Establishing the scale of reduction across parameters		
duration	F1	F2
142	370.7	2508.23
30.60119	464.64	1940.389
111.3988	**−93.94**	**567.8405**

Next, the distances between tokens 2, 3, and 4 and schwa are calculated in step (d).

Table 5d. An example calculation of reduction degree (for Speaker 5). Step (d).

(d) Calculating the distances from averaged tokens to schwa					
Token 2	*Duration*	*F1*	*F2*	Token 3	*Duration*
	109	473.05	2360.85		100.1429
	30.60119	464.64	1940.389		30.60119
	78.39881	**8.41**	**420.4605**		**69.54167**

Step (e) outlines the normalization procedure.

Table 5e. An example calculation of reduction degree (for Speaker 5). Step (e).

(e) Normalizing the distances to the scales across parameters					
Token 2	*Duration*	*F1*	*F2*	Token 3	*Duration*
	78.39881	8.41	420.4605		69.54167
	111.3988	93.94	567.8405		111.3988
	0.703767	**0.089525**	**0.740455**		**0.624259**

Normalization allows one to calculate, as step (f), the degrees of reduction separately for each parameter.

Table 5f. An example calculation of reduction degree (for Speaker 5). Step (f).

(f) Calculating the degrees of reduction across parameters					
Token 2	*Duration*	*F1*	*F2*	Token 3	*Duration*
	1	1	1		1
	0.703767	0.089525	0.740455		0.624259
	0.296233	**0.910475**	**0.259545**		**0.375741**

Finally, step (g) demonstrates how the degree is averaged across parameters under the assumption that duration contributes 50 per cent to the overall degree of reduction, whereas F1 and F2 contribute 25 per cent each (see above for discussion).

Table 5g. An example calculation of reduction degree (for Speaker 5). Step (g).

	token 2	token 3	token 4
F1	0.910475	0.61664	0.949603
F2	0.259545	0.353444	0.386288
	0.58501	**0.485042**	**0.667946**
Duration	0.296233	0.375741	0.325728
F1/F2	0.58501	0.485042	0.667946
	0.440621	**0.430392**	**0.496837**
Token 1	**Token 2**	**Token 3**	**Token 4**
0	**0.440621**	**0.430392**	**0.496837**

The degree of reduction obtained thus for four tokens is presented graphically in Figure 13.

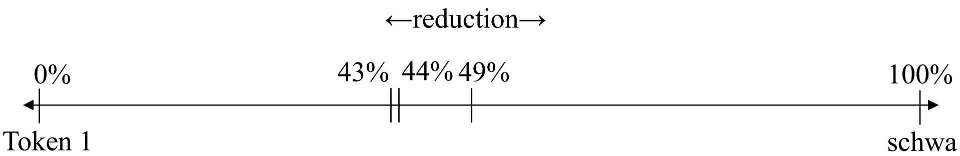

Figure 13. An example of a gradual scale of reduction (for speaker 5).

Because the CPA method is sensitive to inter-speaker differences and normalizes the results obtained, degrees of reduction of /iː/ can now be compared across speakers.

Having outlined CPA for the /iː/ vowel, I suggest that it can also be extended to consonants. A step towards establishing a measure of consonants has been already made by Villafaña Dalcher (2007: 857) who claims that

> lenition has been described in various ways but broadly involves multiple acoustic parameters. These include closure duration, VOT duration, periodicity, intensity, and absence of release burst […]. Because these concrete acoustic characteristics are variable and sometimes independent of one another, it is difficult to quantify the more abstract construct of lenition by referring to its separate components.

My own suggestion includes some of the above parameters, with regard to relevant consonant class. For instance, only plosives can use burst, closure and VOT, which cannot be applied to sonorants for obvious articulatory reasons. As far as consonants are concerned, most casual speech-oriented research just uses duration (e.g. Lavoie 2002). In the case of sonorants, this is unavoidable, but for obstruents, I suggest one should also include incomplete oral closure and (de)voicing, depending on the context/nature of the sound.

Having outlined CPA for vowels and consonants, I suggest that reduction of words can be expressed as the mean of the individual sounds.

The major advantage of the novel method of measuring reduction in vowels and is twofold: (1) it is based on more parameters of reduction than all traditional methods and is therefore potentially more valid, and (2) it makes it possible to express the degree of reduction on a ratio scale, rendering it possible to apply more powerful statistical methods to compute the correlation with the various determinants than in many other studies in the field (Jurafsky's statistically very sophisticated work being a rare exception). In addition (3), the use of a ratio scale has an edge over raw values of formant frequencies in Hertz.

The main disadvantage is that the CPA assumes an equal contribution of duration, F1 and F2 to reduction in vowels, and of incomplete oral closure, (de-)voicing, and/or duration in the case of consonants. This is essentially an arbitrary decision. For example, one might also defend a scheme whereby duration, F1 and F2 each contribute one third. Alternatively, one might argue that because of the traditional view that F1 is perceptually more salient than F2 (see Section 3.1), F1 should weigh more heavily than F2. But until there is conclusive empirical evidence on the relative salience of the temporal and spatial aspects of reduction, it seems reasonable to treat them as being on a par. And whilst this is arbitrary, so is the established practice to select only one parameter at the at the expense of the other ones. At least, the novel approach proposed here does some justice to the complex phonetic make-up of vowels and consonants.

Due to the preliminary nature of the study, only the tense vowel /iː/ was considered and quantified under the CPA approach, whereas other vowels and consonants, for which the CPA has been outlined above, will require follow-up research.

6. Conclusions

The results obtained in the study justify a number of conclusions. Firstly, the parameters of vowel reduction, typically used in an interchangeable fashion in the broad literature in the field, i.e. duration and formant frequencies, fail to exhibit consistent changes in reduction of vowels. Even though this conclusion was validated by a high degree of statistical significance, on a more critical note, they are drawn from a relatively small sample of read speech and so should be approached with some caution.

Despite the preliminary character of the study, several promising directions for further research can be pointed out. First, the linearity hypothesis should be empirically tested against a large corpus of fully spontaneous speech. Second, significant weaknesses were revealed in previous work on reduction, which simplistically treats the phenomenon as an ordinal variable and is often based on crude, auditory impressions.

The study has outlined a novel way to calculate and establish a gradual scale of vowel reduction, based on acoustic analysis and considering all parameters of reduction. I have also showed how the combined parameters approach (CPA) can be extended to consonants in future work, thus promising a comprehensive way to quantify reduction of entire words as the mean of reduction of their individual sounds.

7. Acknowledgements

I would like to express my gratitude to Geoff Schwartz for his most insightful comments on an earlier version of the paper and for kindly sharing his encyclopaedic knowledge of acoustic phonetics. For the great input and extremely useful comments on a research project that involved CPA, I am indebted to Paul Kerswill. I also would like to deeply thank Dafydd Gibbon for his willingness to understand CPA and for his encouragement to simplify its operation. And, of course, I owe very special thanks to Willem Hollmann whose numerous insights led to substantial improvements of both form and substance. But most of all, the paper benefited from a number of very useful comments provided by two anonymous reviewers, contributing to the clarity of presentation of the present paper. I am grateful for their encouragement to pursue the idea of CPA in my further work. All remaining errors are my own.

REFERENCES

Amir, N. and O. Amir. 2007. "Novel measures for vowel reduction". *Proceedings of the 16th International Congress of Phonetic Sciences*, Saarbrücken. 849–852.

Anderson, A. and B. Horwath. 2002. "Referential form and word medial duration in video-mediated and face-to-face dialogues". In: Bos, J., M.E. Foster, and C. Matheson (eds.), *Proceedings of the Sixth Workshop on the Semantics and Pragmatics of Dialogue (EDILOG 2002)*, Edinburgh. Edinburgh: Cognitive Science Centre, University of Edinburgh. 13–20.

Aylett, M. 2001. "Modelling care of articulation with HMMs is dangerous". *EUROSPEECH-2001*. 1491–1494.

Aylett, M. and A. Turk. 2006. "Language redundancy predicts syllabic duration and the spectral characteristics of vocalic syllable nuclei". *Journal of the Acoustical Society of America* 119. 3048–3058.

Aylett, M. and A. Turk. 2004. "The Smooth Signal Redundancy Hypothesis: A functional explanation for relationships between redundancy, prosodic prominence and duration in spontaneous speech". *Language and Speech* 47(1). 31–56.

Aylett, M. and A. Turk. 1998. "Vowel quality in spontaneous speech: What makes a good vowel?" *Proceedings of the 5th International Conference on Spoken Language Processing, Incorporating the 7th Australian International Speech Science and Technology Conference*, Sydney. Sydney: Australian Speech Science and Technology Association (ASSTA).

Baker, R.E. and A.R. Bradlow. 2009. "Variability in word duration as a function of probability, speech style and prosody". *Language and Speech* 52. 391–413.

Barry, W. and B. Andreeva. 2001. "Cross-language similarities and differences in spontaneous speech patterns". *Journal of the International Phonetic Association* 31. 51–66.
Beckford Wassink, A., R.A. Wright and A.D. Franklin. 2007. "Intraspeaker variability in vowel production: On investigation of motherese, hyperspeech and Lombard speech in Jamaican speakers". *Journal of Phonetics* 35. 363–379.
Bell, A., J. Brenier, M. Gregory, C. Girand and D. Jurafsky. "Predictability effects on durations of content and function words in conversational English". *Journal of Memory and Language* 60(1). 92–111.
Bell, A., D. Jurafsky, E. Fossler-Lussier, C. Girand, M. Gregory and D. Gildea. 2003. "Effects of dislfuencies, predictability and utterance position on word form variation in English conversation". *Journal of the Acoustical Society of America* 113. 1001–1024.
Bondarko, L.V., N.B. Volskaya, S.O. Tananaiko and L.A. Vasilieva. 2003. "Phonetic properties of Russian spontaneous speech". *Proceedings of the 15th International Congress of Phonetic Sciences*, Barcelona. 2973–2976.
Bradlow, A.R., G.M. Torretta and D.B. Pisoni. 1996. "Intelligibility of normal speech. Global and fine-grained acoustic-phonetic talker characteristics". *Speech Communication* 20. 255–272.
Bybee, J. 2000. "Lexicalization of sound change and alternating environments". In: Broe, M. and J. Pierrehumbert (eds.), *Papers in laboratory phonology V*. Cambridge: Cambridge University Press. 250–269.
Bybee, J. and J. Scheibman.1999. "The effect of usage on degrees of constituency: The reduction of *don't* in English". *Linguistics* 37. 575–596.
Byrd, D. and C.C. Tan. 1996. "Saying consonants clusters quickly". *Journal of Phonetics* 24. 263–282.
Carlson, R., G. Fant and B. Granström. 1975. "Two-formant models, pitch and vowel perception". In: Fant, G. and M.A.A. Tatham (eds.), *Auditory analysis and perception of speech*. London: Academic Press. 55–82.
Clark, L. and G. Trousdale. 2009. "Exploring the role of token frequency in phonological change: Evidence from TH-Fronting in east-central Scotland". *English Language and Linguistics* 13. 33–55.
Corbin, O. 2003. "Phoneme deletion in spontaneous British English". *Proceedings of the 15th International Congress of Phonetic Sciences*, Barcelona. 2813–2816.
Cruttenden, A. 2008. *Gimson's pronunciation of English*. (7th edition.) London: Hodder.
De Jong, G., K. McDougall, T. Hudson and F. Nolan. 2007. "The speaker discriminating power of sounds undergoing historical change: A formant-based study". *Proceedings of the 16th International Congress of Phonetic Sciences*, Saarbrücken, August 2007. 1813–1816.
Delattre, P. 1965/1981. "An acoustic and articulatory study of vowel reduction in four languages". In: Malmberg, B.(ed.), *Studies in comparative phonetics*. Heidelberg: Julius Groos Verlag. 63–94.
Duez, D. 2001. "Restoration of deleted and assimilated consonant sequences in conversational French: Effects of preceding and following context". *Journal of the International Phonetic Association* 31. 101–114.
Ellis, L. and W. Hardcastle. 1999. "An instrumental study of alveolar to velar assimilation in careful and fast speech". *Proceedings of the 14th International Congress of Phonetic Sciences*, San Francisco. 2425–2428.
Ellis, L. and W. Hardcastle. 2002. "Categorical and gradient properties of assimilation in alveolar to velar sequences: Evidence from EPG and EMA". *Journal of Phonetics* 30. 373–396.
Engstrand, O. and D. Krull. 2002. "Duration of syllable-sized units in casual and elaborated speech: Cross-language observations on Swedish and Spanish". *Proceedings from Fonetik*

2002, the 14th Swedish Phonetics Conference, Stockholm. (*Speech, Music and Hearing, Quarterly Progress and Status Report (TMH-QPSR)* 44.) 69–72.

Fant, G., U. Stalhammer and I. Karlsson. 1974. "Swedish vowels in speech materials of various complexity". *Speech Communication* 2. 139–147.

Ferragne, E. and F. Pellegrino. 2010. "Formant frequencies of vowels in 13 accents of the British Isles". *Journal of the International Phonetic Association* 40. 1–34.

Fowler, C. and J. Housum.1987. "Talkers signalling of new and old words in speech and listeners perception and use of the distinction". *Journal of Memory and Language* 26. 489–504.

Greenberg, S. and E. Fossler-Lussier. 2000. "The uninvited guest: Information's role in guiding the production of spontaneous speech". *Proceedings of the CREST workshop on models of speech production: Motor planning and articulatory modeling.* Kloster Seeon, Germany. 1–4.

Harris, J. 2005. "Vowel reduction as information loss". In: Carr, P., J. Durand and C.J. Ewen (eds.), *Headhood, elements, specification and contrastivity*. Amsterdam: Benjamins. 119–132.

Hillenbrand, J.M., L.A. Getty, M.J. Clark and K. Wheeler. 1995. "Acoustic characteristics of American English vowels". *Journal of the Acoustical Society of America* 97. 3099–3111.

Hollmann, W.B. and A. Siewierska. 2006. "Corpora and (the need for) other methods in a study of Lancashire dialect". *Zeitschrift für Anglistik und Amerikanistik* 54. 203–216.

Hollmann, W.B. and A. Siewierska. 2007. "A construction grammar account of possessive constructions in Lancashire dialect: some advantages and challenges". *English Language and Linguistics* 11. 407–424.

Howes, D. 1967. "Equilibrium theory of word frequency distributions". *Psychonometrics Bulletin* 1. 1–18.

Jaworski, S. 2008. Lenition processes in English and other languages: A hierarchy of susceptibility to inertia. (Unpublished PhD dissertation, Adam Mickiewicz University, Poznań.)

Joos, M. 1948. "Acoustic phonetics". *Language* 24. 1–136.

Jurafsky, D., A. Bell, E. Fossler-Lusier, C. Girana and W. Raymond.1998. "Reduction of English Function words in SWITCHBOARD". *Proceedings of ICLSP98*. 3111–3114.

Kelly, J. and J. Local. 1989. *Doing phonology*. Manchester: Manchester University Press.

Koopmans-van Beinum, F.J. 1980. Vowel contrast reduction: An acoustic and perceptual study of Dutch vowels in various speech conditions. (Unpublished PhD dissertation, University of Amsterdam.)

Lavoie, L. 2002. "Some influences on the realization of for and four in American English". *Journal of the International Phonetic Association* 32. 175–202.

Lennes, M, N. Alarotu and M. Vainio. 2001. "Is the phonetic quality of unaccented vowels predictable? An example from spontaneous Finnish". *Journal of the International Phonetic Association* 31. 127–138.

Lindblom, B. 1990. "Explaining phonetic variation: A sketch of the H&H theory". In: Hardcastle, W.J. and A. Marchal, A. (eds.), *Speech production and speech modeling*. Dordrecht: Kluwer Academic Publishers. 403–439.

Lindblom, B. 1963. "Spectrographic study of vowel reduction". *Journal of the Acoustical Society of America* 35. 1773–1781.

Local, J. 2003. "Variable domains and variable relevance: Interpreting phonetic exponents". *Journal of Phonetics* 31. 321–339.

Lodge, K. 1984. *Studies in the phonology of colloquial English*. London: Croom Helm.

Moon, S.J. and B. Lindblom. 1994. "Interaction between duration, context, and speaking style in English stressed vowels". *Journal of the Acoustical Society of America* 96: 40–55.

Max, L and A.J. Caruso. 1997. “Acoustic measures of temporal intervals across speaking rates: Variability of syllable- and phrase-level relative timing”. *Journal of Speech, Language and Hearing Research* 40. 1097–1110.

Mitterer, H. and M. Ernestus. 2006. “Listeners recover /t/s that speakers reduce: Evidence from /t/-lenition in Dutch”. *Journal of Phonetics* 34. 73–103.

Moon, S.J. and B. Lindblom. 1994. “Interaction between duration, context, and speaking style in English stressed vowels”. *Journal of the Acoustical Society of America* 96. 40–55.

Mooshammer, C. and C. Geng. 2008. “Acoustic and articulatory manifestations of vowel reduction in German”. *Journal of the International Phonetic Association* 38(2). 119–136.

Munson, B. and N.P. Solomon. “The effect of phonological neighborhood density on vowel articulation”. *Journal of Speech, Language, and Hearing Research* 47. 1048–1105.

Newman, R.S. and S. Evers. 2007. “The effect of talker familiarity on stream segregation”. *Journal of Phonetics* 35. 85–103.

Ohala, J. and H. Kawasaki. 1984. “Prosodic phonology and phonetics”. *Phonology* 1. 113–128.

Oktaba, W. 1966. *Elementy statystyki matematycznej i metodyka doświadczalnictwa* [Elements of mathematical statistics and research methodology]. Warszawa: PWN.

Picheny, M., N. Durlach and L.Braida. 1986. “Speaking clearly for the hard of hearing. Acoustic characteristics of clear and conversational speech”. *Journal of Speech and Hearing Research* 29. 434–445.

Piroth, H.G. and P.M. Janker. 2004. “Speaker-dependent differences in voicing and devoicing of German obstruents”. *Journal of Phonetics* 32. 81–109.

Port, R.F. 1981. “Linguistic timing factors in combination”. *Journal of the Acoustical Society of America* 69. 262–274.

Recasens, D. 2004. “The effect of syllable position on consonant reduction: Evidence from Catalan consonant clusters”. *Journal of Phonetics* 32. 435–453.

Rosenhouse, J. and L. Kishon-Rabin. 2003. “The effect of speech rate and noise on bilinguals' speech perception: the case of native speakers of Arabic in Israel”. *EUROSPEECH-2003*. 2089–2092.

Saltzman, E. 2003. “Temporal aspects of articulatory control”. *EUROSPEECH-2003*. 2565–2568.

Sanford, D. 2008. “Metaphor and phonological reduction in English idiomatic expressions”. *Cognitive Linguistics* 19. 585–603.

Shearme, J.N. and J.N. Holmes. 1962. “An experimental study of the classification of speech sounds according to their distribution in the formant 1–formant 2 plane”. *Proceedings of the 4th International Congress of Phonetic Sciences*, Helsinki. 234–240.

Shockey, L. 2003. *Sound patterns of spoken English*. Oxford: Blackwell.

Shockey, L. 1974. “Phonetic and phonological properties of connected speech”. *Ohio State University Working Papers in Linguistics* 17. iv–143.

Silva, V. de, A. Iivonen, L.V. Bondarko and L.C.W. Pols. 2003. “Common and language dependent phonetic differences between read and spontaneous speech in Russian, Finnish and Dutch”. *Proceedings of the 15th International Congress of Phonetic Sciences*, Barcelona. 2977–2980.

Summers, W.V. 1987. “Effects of stress and final-consonant voicing on vowel production: Articulatory and acoustic analyses”. *Journal of the Acoustical Society of America* 82(3). 847–863.

Tagliamonte, S.A and R.V. Roeder. 2009. “Variation in the English definite article: Socio-historical linguistics in t'speech community”. *Journal of Sociolinguistics* 13(4). 435–471.

Tiffany, W. R. 1959. “Non random sources of variation in vowel quality”. *Journal of Speech and Hearing Research* 2. 305–317.

Van Bergem, D.R. 1993. "Acoustic vowel reduction as a function of sentence accent, word stress, and word class". *Speech Communication* 12. 1–23.
Villafaña Dalcher, C. 2007. "Statistical methods for quantitative analysis of multiple lenition components". *Proceedings of the 16th International Congress of Phonetic Sciences*, Saarbrücken. 857–860.
Wright, R. 2003. "Factors of lexical competition in vowel articulation". In: Local, J., R. Ogden and R. Temple, R. (eds.), *Papers in Laboratory Phonology V*. Cambridge: Cambridge University Press. 75–87.
Zhao, Y. and D. Jurafsky. 2009. "The effect of lexical frequency and Lombard reflex on tone hyperarticulation". *Journal of Phonetics* 27. 231–247.

Address correspondence to:
Małgorzata Kul
School of English
Adam Mickiewicz University
Al. Niepodległości 4
61-874 Poznań
Poland
kgosia@ifa.amu.edu.pl