Journal of Phonetics 49 (2015) 1–15

Contents lists available at ScienceDirect

Journal of Phonetics

journal homepage: www.elsevier.com/locate/phonetics

Research Article

Regional variation in temporal organization in American English

Cynthia G. Clopper [a,*], Rajka Smiljanic [b]

[a] Department of Linguistics, Ohio State University, 1712 Neil Avenue, 222 Oxley Hall, Columbus, OH 43210, USA
[b] Department of Linguistics, University of Texas at Austin, 305 E. 23rd Street, Austin, TX 78712, USA

A R T I C L E I N F O

Article history:
Received 2 July 2013
Received in revised form
2 October 2014
Accepted 9 October 2014
Available online 18 November 2014

Keywords:
Temporal organization
Rhythm
Dialect variation
American English

A B S T R A C T

The goal of the current study was to explore the temporal organization of six regional dialects of American English to gain a better understanding of the perceptual impressions of speaking rate variation. The study further examines whether regional dialects form different groupings based on their segmental vs. global temporal characteristics. Acoustic measures included articulation rate, pause frequency and duration, and vowel and consonant duration variability. The results revealed that Southern American English is characterized by a slow overall articulation rate, long pauses, and highly variable syllable-to-syllable vowel durations, whereas the New England dialect is characterized by a fast overall articulation rate, short pauses, and highly variable syllable-to-syllable consonant durations. The patterns for the other dialects are more mixed: the Northern and Western dialects are characterized by low variability vowel durations, the Midland dialect shares a slow articulation rate with the Southern dialect, and the Mid-Atlantic dialect exhibits no unique temporal properties among those examined. Thus, temporal variation in regional dialects of American English is orthogonal to vowel variation, in which New England, Midland, and Western dialects are often characterized together as “General American”. Taken together, the results are consistent with the stereotype that Southerners talk slowly and Northerners talk quickly and suggest that pausing and segmental duration variability may contribute to the perceived speaking rate differences.

1. Introduction

One of the many stereotypes about regional varieties of American English is that Southerners talk slowly and Northerners, particularly Northeasterners, talk quickly (Niedzielski & Preston, 2003). However, a deeper understanding of what contributes to these perceptions is lacking and the empirical sociophonetic research on one aspect of temporal variation in American English, namely speaking rate differences, is somewhat mixed. Although Jacewicz and colleagues (Jacewicz, Fox, O'Neill, & Salmons, 2009; Jacewicz, Fox, & Wei, 2010) observed faster speaking rates among Wisconsin talkers than North Carolina talkers, consistent with this stereotype, Ray and Zahn (1990), Byrd (1994), and Clopper and Smiljanic (2011) all reported marginal or null results in comparisons between Southern and Northern or Midwestern varieties. However, Byrd (1994) and Clopper and Smiljanic (2011) both observed more frequent pausing among Southern talkers than talkers from other regions and Jacewicz et al. (2010) and Kendall (2009) reported longer pauses among Southern talkers than Northern or Midland talkers, respectively. Focusing on segmental duration variation across regional varieties, Clopper, Pisoni, and de Jong (2005) found that Southern lax vowels are longer than Northeastern and Western lax vowels, but that tense vowels do not differ in duration across dialects. Similarly, Jacewicz, Salmons, and Fox (2007) found that Southern front vowels are longer than Northern front vowels. Taken together, these results suggest that the speaking rate stereotype may be based on factors other than overall articulation rate, including pause frequency, pause duration, and vowel duration.

Recently, White, Mattys, and Wiget (2012) explored listener sensitivity to durational cues in differentiating between different languages and between varieties of the same language. They used “bleached” sentences in which information about specific segments is removed but the overall durational patterns are retained (following Ramus, Dupoux, & Mehler, 2003). In White et al.'s (2012) study, English listeners were predictably able to differentiate between English and Spanish, which are known to exhibit

* Corresponding author.
E-mail addresses: clopper.1@osu.edu (C.G. Clopper), rajka@austin.utexas.edu (R. Smiljanic).

0095-4470/$ - see front matter © 2014 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.wocn.2014.10.002

substantial timing differences. Surprisingly, however, they were also able to differentiate among varieties of British English, suggesting that a host of durational cues, including speech rate, as well as durational contrasts and utterance-final lengthening, plays a role in categorization between and within languages. Given the evidence that variation in temporal patterns contributes to the differentiations among varieties of the same language, the goal of the current study was to explore in detail variability in temporal organization, beyond speaking rate, including pausing and variability in consonant and vowel interval durations, in six regional varieties of American English. The results will enhance our understanding of speech properties that underlie differentiation of regional American English varieties beyond typically-examined vowel characteristics (e.g., Clopper et al., 2005; Labov, Ash, & Boberg, 2006; Thomas, 2001) and provide further support for the empirical basis of the widely held social stereotypes about regional differences in speaking rate. This study also sheds new light on the relationships of similarities and differences among these American English varieties. A direct comparison of six regional varieties expands on previous studies which have typically examined only two or three dialects (cf., Byrd, 1994). In addition, the comprehensive analysis of vowel and consonant duration variability, as a metric of temporal organization, provides a novel perspective on regional temporal organization variation in American English.

Variation in vowel and syllable duration has been examined for a number of national, regional, and ethnic varieties of English and the observed differences are typically attributed to contact with languages with less syllable-to-syllable vowel variation. For example, British English exhibits more variation in vowel duration than Singapore English (Deterding, 2001; Low, Grabe, & Nolan, 2000) or New Zealand English (Warren, 1998), presumably due to ongoing contact with Mandarin in Singapore and historical contact with Maori in New Zealand. Comparisons between Maori English and Pakeha English (Pakeha is used to refer to New Zealanders of European descent) in New Zealand provide further evidence for the role of language contact in this variation. Maori English exhibits less vowel reduction than Pakeha English (Holmes & Ainsworth, 1996, 1997) and therefore less vowel and syllable duration variability (Szakay, 2006; Warren, 1998). Pakeha English also exhibits a reduction in syllable duration variation over time (Nokes & Hay, 2012), suggesting that the temporal organization of New Zealand English is changing as a result of contact with Maori and Maori English. Language contact may also explain the variability observed in temporal organization in ethnic varieties of North American English. Jamaican English, Hispanic English, Native American English (Cherokee and Lumbee), and Chinese-American English have all been described as exhibiting less variability in syllable duration than white American English as a result of influences from substrate or heritage languages (Coggshall, 2008; Newman & Wu, 2011; Thomas & Carter, 2006). Similar variation has been observed in British English, in which urban varieties spoken by ethnically heterogeneous populations exhibit less vowel duration variability than suburban varieties spoken by homogeneous British populations (Torgersen & Szakay, 2012). White and Mattys (2007) have also demonstrated that regional varieties of British English differ in the degree of reduction of unstressed syllables and that these differences contribute to variation in temporal organization.

The metrics used in these previous studies to explore variability in vowel and syllable duration were originally developed to test the purported rhythmic distinction between stress-timed and syllable-timed languages (e.g., Abercrombie, 1967). According to this hypothesis, syllable-timed languages, such as Spanish, have isochronous syllables and stress-timed languages, such as English, have isochronous stressed syllables. Furthermore, stress-timed languages typically exhibit vowel reduction processes and a large number of syllable types, whereas syllable-timed languages typically exhibit simpler syllable types and little vowel reduction. Syllable, vowel, and consonant durations should therefore be less variable in a syllable-timed language than a stress-timed language. A number of metrics have been proposed to capture this difference in segment and syllable variability. Ramus, Nespor, and Mehler (1999) proposed $\%V$ (the proportion of vocalic intervals), ΔV (the standard deviation of vocalic interval duration), and ΔC (the standard deviation of consonant interval duration) and observed a negative correlation between $\%V$ and ΔC across languages. Languages with a relatively large proportion of vocalic segments (high $\%V$) and relatively little variability in consonant interval duration (low ΔC) were identified as syllable-timed and languages with a relatively low $\%V$ and a relatively high ΔC were identified as stress-timed. Dellwo and Wagner (2003) noted that the ΔC and ΔV metrics were strongly correlated with speaking rate and Dellwo (2006) therefore proposed the *VarcoC* and *VarcoV* metrics, which normalize ΔC and ΔV, respectively, for speaking rate. Grabe and Low (2002) proposed an additional metric, the pairwise variability index (PVI), to capture the greater variability in syllable duration in stress-timed languages than syllable-timed languages. The PVI is the average relative duration of consecutive syllables and a high PVI therefore corresponds to regular alternations between long and short syllables. Although languages like English and German do not exhibit strictly alternating long and short syllables, they do exhibit substantial vowel-to-vowel duration variability due to lexical stress and phonological vowel length distinctions between tense and lax vowels (English), long and short vowels (German), and monophthongs and diphthongs (English and German). English and German therefore have higher vowel PVI values than languages like French and Spanish which exhibit less vowel-to-vowel duration variability. Grabe and Low (2002) further proposed normalizing the vowel PVI metric for speaking rate, but argued that the raw consonant PVI metric was sufficient to capture rhythmic variability across languages.

Although most of the previous research documenting within-language variability in temporal organization has relied on these metrics for analysis, several recent studies have raised critical questions about their reliability and validity. In particular, the metrics exhibit substantial variability within languages when different speaking styles (e.g., read vs. conversational speech) and different talkers are compared (Arvaniti, 2009, 2012; Loukina, Kochanski, Rosner, Keane, & Shih, 2011; Nolan & Asu, 2009; Smiljanic & Bradlow, 2008; Wiget, White, Schuppler, Grenon, Rauch, & Mattys, 2010). More fundamentally, however, the different metrics often point to different interpretations of stress- vs. syllable-timing because they can also capture differences in vowel reduction, vowel length contrasts, resolution of vowel–vowel hiatus, and consonant cluster simplification (Dauer, 1983; Fagyal, 2010). This latter kind of variation is exactly how we expect regional dialects of American English to vary, however, so we used these metrics to explore this aspect of variation in temporal organization in the current study.

To reduce the effects of speaking style and individual talker variability on our analysis, we obtained our measurements from two read passages produced by each of 60 talkers, for a total of more than 1 h of speech. Read passages were selected instead of isolated sentences or interview speech because they exhibit two desirable characteristics for this type of analysis. First, the use of read passages ensures that the same linguistic content was produced by all of the talkers, eliminating a potential confound between regional dialect and phonological content, such as syllable complexity or vowel token frequency. Given the substantial reported effect of linguistic content on the various timing measures (Wiget et al., 2010), this control over the content was essential for producing interpretable results. Second, the use of read passages involves long stretches of connected speech, which allows for an analysis of pausing phenomena. Read passages therefore provide a reasonable compromise between isolated read sentences, in which control over the content is maximized, and interview speech, in which naturally occurring temporal variation is maximized. This intermediate status of read passages is further supported by Arvaniti's (2012) finding that read passages provide an estimate of timing that falls between the estimates obtained from isolated sentences and interview speech. Any findings in this study regarding variation in the temporal properties of read speech may, therefore, be more exaggerated in spontaneous speech. Read speech may also reflect individual talker differences in style and/or reading ability, especially among children and older adults (Jacewicz et al., 2010). Although these sources of variability may lead to different patterns of variation in the temporal organization of read speech relative to spontaneous speech, we expect that this potential variability will be minimized in our sample of college-educated young adults, who are skilled and fluent readers (see Jacewicz et al., 2010). The total amount of speech that was analyzed per dialect (approximately 10 min) is comparable to the amount of read speech that was analyzed per speaking style by Smiljanic and Bradlow (2008) and per language by Arvaniti (2012), and is substantially more than the amount of speech that was analyzed in most previous studies, which ranges from as few as 30 sentences (White & Mattys, 2007) to as many as 100 sentences (Low et al., 2000) per variety.

2. Methods

2.1. Materials

The materials consisted of two read passages produced by five male and five female talkers from each of six regional dialects of American English (New England, Mid-Atlantic, North, Midland, South, and West) from the Nationwide Speech Project corpus (Clopper & Pisoni, 2006). The two passages were the first paragraph of the Rainbow Passage (Fairbanks, 1940) and the beginning of the Goldilocks fairytale (Stockwell, 2002). The 60 talkers were all monolingual native speakers of American English who had lived exclusively in their dialect region until at least the age of 18 years old. Both parents of each talker were also raised in the same region. Although some varieties of New England and Southern American English are non-rhotic, all of the talkers in the NSP corpus are rhotic. The passages were recorded as part of a recording session that also included read words and sentences and an interview with the first author. The talkers were 18–25 years old and enrolled as students at Indiana University at the time of the recording.

2.2. Acoustic measurements

Two sets of measurements were obtained from each passage. The first set of measurements was selected to explore variation across regional dialects in speaking rate and included articulation rate, pause frequency, and pause duration. Each passage was first coded by hand for pauses, disfluencies, and recording errors. Pauses were defined as periods of silence longer than 100 ms, based on Wennerstrom and Siegel's (2003) analysis of pause duration in conversational speech (see also Clopper & Smiljanic, 2011; Henderson, Goldman-Eisler, & Skarbek, 1966). Disfluencies, including hesitations, backtracking, and false starts, and technical recording errors were excluded from the analysis. Disfluencies comprised less than 2% of the total speech produced in the passages, suggesting relatively high reading proficiency among our college-educated talkers. Articulation rate was defined over inter-pause intervals as the number of syllables per second of fluent, connected speech. Syllable counts were based on the underlying form of the word sequences produced by each talker, including lexical and/or morphological insertions, deletions, and substitutions relative to the provided text. Thus, the syllable counts varied modestly across talkers as a result of minor errors in reading the text. Pause frequency was defined as the number of pauses longer than 100 ms per passage for each talker. Pause duration was defined as the mean duration of all pauses longer than 100 ms in each passage for each talker.

The second set of measurements was selected to capture variation in the relative duration of consonant and vowel intervals across dialects and included %*V*, Δ*C*, ΔV, *VarcoC*, *VarcoV*, *rPVI-C*, and *nPVI-V*. Following Grabe and Low (2002) and to permit comparison with previous research (e.g., Arvaniti, 2009, 2012), we have included the normalized PVI measure for vowels, but the unnormalized PVI measure for consonants. In addition, although the *VarcoV* and *nPVI-V* measures were both developed to capture rate-normalized vowel duration variability and the Δ*C* and *rPVI-C* measures were both developed to capture unnormalized consonant duration variability, we have included all three sets of measures in our analysis to allow for more direct comparisons between our study and the previous literature in which various subsets of these measures are reported.

To obtain these measures, consonant and vowel intervals were transcribed by trained research assistants following the transcription conventions described by Ramus et al. (1999). Specifically, sequences of consonants or vowels (within and across syllable and word boundaries) were transcribed as a single C or V interval, respectively. Consonants and vowels were defined phonologically, with the exception of coda approximants, which were treated as vowels. Coda approximants that were resyllabified across a word boundary as an onset were treated as consonants. Fig. 1 shows example transcriptions of the phrases, “a daddy bear, a mommy bear” from the Goldilocks passage produced by a New England female talker (top) and a Southern male talker (bottom).

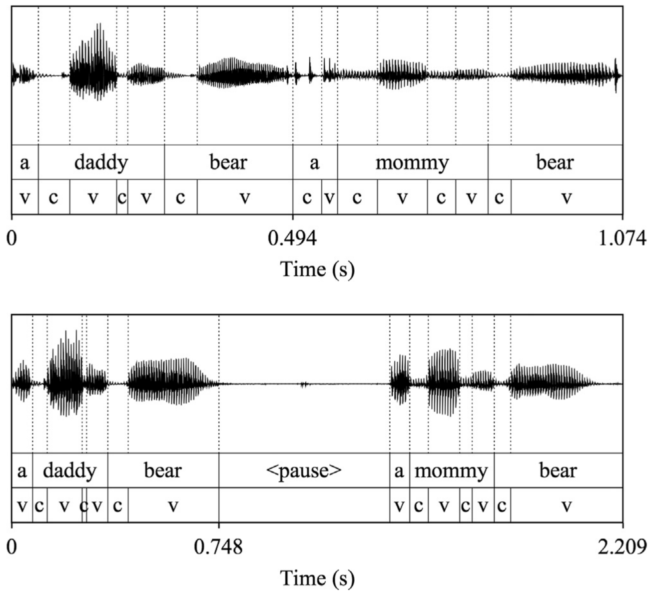

Fig. 1. Example transcriptions from the Goldilocks passage of the phrases “a daddy bear, a mommy bear” produced by a New England female talker (top) and a Southern male talker (bottom).

Note that the word *bear* is always transcribed as a CV, with the coda approximant included in a single V interval with the nucleus vowel. Further, the second *a* produced by the New England female talker is transcribed as a CV because it was realized with a glottal stop onset; the coda approximant in *bear* was not resyllabified as an onset.

All transcriptions were hand-checked by the first author and the durations of the consonant and vowel intervals were then extracted. The complete data set included 1658 inter-pause intervals comprising 36,846 consonant and vowel intervals. The five metrics were calculated separately for each inter-pause interval: %*V* was defined as the total duration of the vowel intervals divided by the total duration of the inter-pause interval, ΔC and ΔV were defined as the standard deviation of the durations of consonant and vowel intervals, respectively, *VarcoC* and *VarcoV* were defined as the standard deviation divided by the mean duration of the consonant and vowel intervals, respectively, *rPVI-C* was defined as the average pairwise absolute difference between consecutive consonant intervals, and *nPVI-V* was defined as the average pairwise absolute difference between consecutive vowel intervals divided by their mean duration. The dependent measures were all reasonably normally distributed and were therefore analyzed without transformation.

2.3. Statistical analysis

The two pause measures were calculated separately for each talker for each passage and the speaking rate measure and the seven consonant and vowel interval measures were calculated separately for each inter-pause interval for each talker. The calculations of the consonant and vowel variability measures (ΔC, ΔV, *VarcoC*, *VarcoV*, *rPVI-C*, and *nPVI-V*) require at least two intervals of the same type (C or V). Inter-pause intervals containing fewer than four total consonant and vowel intervals were therefore excluded from the analysis (0.8% of the data) and each inter-pause interval included in the analysis consisted of at least a CVCV or VCVC sequence. Each of the dependent variables was analyzed using a linear mixed effects model with regional dialect, talker gender, and their interaction as fixed effects. Talker gender was included in the models because previous research has shown that men talk faster and produce longer pauses than women (Byrd, 1994; Jacewicz et al., 2009, 2010; Kendall, 2009; Kowal & O'Connell, 1980; Whiteside, 1996; cf., Ray & Zahn, 1990). Previously reported gender effects for pause frequency are mixed, with some studies showing more frequent pauses for men than women (Kowal & O'Connell, 1980) and other studies showing the opposite pattern (Whiteside, 1996). Thus, gender was expected to be a significant factor in the measures examined in this study. Log-likelihood comparisons were used to determine if talker gender and/or the gender $\times$ dialect interaction were significant for each of the dependent variables.

A number of covariates were also included in the analyses. First, given that speaking rate has been shown to correlate with pause frequency and with consonant and vowel interval metrics (Dellwo, 2006; Dellwo & Wagner, 2003; Jacewicz et al., 2010), articulation rate was included as a covariate fixed effect in these analyses. Given that Kendall (2009) found that read speech is produced with a slower speaking rate and more pauses than interview speech, the inclusion of speaking rate as a covariate in our models may allow for greater generalization from read speech to other registers. Second, Kendall (2013) found that speaking rate increased throughout the read passages in his analysis. We therefore included the inter-pause interval (numbered consecutively from the beginning of the passage) as a covariate in the analysis of articulation rate. Finally, to explore the relationship between pause duration and pause frequency, pause duration was included as a covariate in the analysis of pause frequency and vice versa.

For all models, female was the reference level of the talker gender fixed effect. The reference levels for regional dialect were selected separately for each dependent variable so that comparisons were made first to the dialect with the lowest mean value and

Table 1
Mean articulation rate, mean number of pauses per passage, and mean pause duration for each talker dialect. Number of pauses is further divided by talker gender.

Dialect	Articulation rate (syll/s)	Number of pauses		Pause duration (ms)
		Females	Males	
New England	5.73	11.6	13.4	422
Mid-Atlantic	5.62	11.9	12.4	456
North	5.54	13.0	12.6	462
Midland	5.36	14.6	11.6	437
South	5.40	10.8	16.6	487
West	5.50	12.7	14.0	439

then, in a second model, to the dialect with the highest mean value. To account for subject and item variability, random intercepts for talker, passage, and their interaction were examined in each model.[1] A random talker slope for articulation rate was also considered in each of the models containing articulation rate as a covariate fixed effect. Log-likelihood comparisons confirmed that the random intercept for the talker × passage interaction and the random talker slope for articulation rate were not significant in any of the models. The random intercepts for talker and passage were retained in all models to allow for a consideration of talker and passage variance. Markov chain Monte Carlo simulations were used to estimate *p*-values for all models (Baayen, Davidson, & Bates, 2008). Full model outputs for each analysis are provided in the Appendix.

3. Results

3.1. Speaking rate

A summary of the mean articulation rate, pause frequency, and pause duration for each talker dialect is shown in Table 1. The overall mean articulation rate across all dialects was 5.53 syllables per second (SD = 0.45) and the talkers produced an average of 12.9 pauses per passage (SD = 5.3) with a mean pause duration of 450 ms (SD = 71 ms). Although the minimum pause duration was defined as 100 ms, the distribution of pause durations was only moderately positively skewed and the median pause duration was 435 ms.

The linear mixed effects models revealed a significant effect of talker dialect on articulation rate, which was faster for the New England talkers than for the Midland talkers ($\beta = 0.37$, $t = 1.94$, $pMCMC = 0.030$) and, marginally, for the Southern talkers ($\beta = 0.34$, $t = 1.78$, $pMCMC = 0.052$). The effect of gender and the interaction between dialect and gender were not significant. However, inter-pause interval was a significant predictor of articulation rate ($\beta = 0.01$, $t = 2.79$, $pMCMC = 0.021$). As in Kendall's (2013) analysis, articulation rate increased throughout the duration of the passage. The model variance across passages approached zero, suggesting similar within-talker articulation rates across passages. Consistent with stereotypes regarding Northern vs. Southern varieties of American English, the New England talkers spoke the fastest and the Midland and Southern talkers spoke the slowest. The dialect differences were significant even when inter-pause interval was included in the model, confirming that the effect of dialect is independent of passage-internal increases in speaking rate.

Unlike articulation rate, pause frequency varied with both dialect and gender and the mean number of pauses are separated by gender in Table 1. The linear mixed effects models revealed significant main effects of articulation rate ($\beta = -2.64$, $t = -3.85$, $pMCMC < 0.001$) and pause duration ($\beta = -8.21$, $t = -1.87$, $pMCMC = 0.034$), as well as significant differences between male and female Southern talkers ($\beta = 6.00$, $t = 4.17$, $pMCMC < 0.001$), between Midland and Southern female talkers ($\beta = 3.10$, $t = 2.17$, $pMCMC = 0.028$), and between Southern male talkers and New England ($\beta = 4.37$, $t = 2.12$, $pMCMC = 0.020$), Mid-Atlantic ($\beta = 4.64$, $t = 2.32$, $pMCMC = 0.013$), Northern ($\beta = 5.47$, $t = 2.73$, $pMCMC = 0.004$), and Midland ($\beta = 8.38$, $t = 4.19$, $pMCMC < 0.001$) male talkers. The model variance across passages was large relative to the variance across talkers, reflecting the overall difference in passage length. The Goldilocks passage is longer and therefore exhibited more pauses than the Rainbow passage. The relationship between articulation rate and number of pauses is negative, suggesting that talkers who spoke more slowly also paused more frequently. The relationship between pause duration and number of pauses is also negative, suggesting that talkers who paused more frequently produced shorter average pauses that talkers who paused less frequently. Among the female talkers, the Midland talkers paused the most and the Southern talkers paused the least. However, among the male talkers, the Southern talkers paused the most and the Midland talkers paused the least. These results were significant even when articulation rate and pause duration were included in the model, confirming that the effects of dialect and gender on pause frequency are independent of the effect of articulation rate and pause duration.

Regional dialect, talker gender, articulation rate, and pause frequency also contributed to variability in pause duration. The linear mixed effects models revealed significant main effects of articulation rate ($\beta = -48.98$, $t = -3.65$, $pMCMC < 0.001$), pause frequency ($\beta = -2.35$, $t = -2.92$, $pMCMC = 0.047$), and gender ($\beta = 70.13$, $t = 5.89$, $pMCMC < 0.001$), as well as significant differences between

[1] Passage was treated as the random item effect rather than inter-pause interval to avoid over-fitting the data. The 1658 inter-pause intervals included in the analysis consisted of 280 unique strings, of which 148 (53%) occurred only once in the data. The frequency distribution of the unique strings was highly skewed, with a maximum of 55 and a mean of 5.9 tokens per string. Specifying random intercepts for inter-pause intervals would therefore potentially result in near-perfectly fitting over half of the unique strings and over-fitting many of the others.

Table 2
Mean %V, mean ΔC, mean ΔV, mean *VarcoC*, mean *VarcoV*, mean *rPVI-C*, and mean *nPVI-V* for each talker dialect.

Dialect	%*V*	Δ*C* (ms)	Δ*V* (ms)	*VarcoC*	*VarcoV*	*rPVI-C*	*nPVI-V*
New England	49.84	56.74	55.55	59.13	55.66	67.61	58.15
Mid-Atlantic	50.25	56.85	57.77	57.79	56.18	66.86	58.95
North	49.40	56.66	52.97	58.78	53.38	67.08	57.40
Midland	50.27	59.46	59.64	58.32	55.70	70.40	58.72
South	51.54	54.72	61.95	55.64	57.37	65.26	60.85
West	49.26	56.97	53.65	58.06	53.69	67.14	56.62

Southern talkers and New England ($\beta=51.20$, $t=2.44$, $pMCMC=0.008$), Midland ($\beta=52.53$, $t=2.56$, $pMCMC=0.004$), and Western ($\beta=42.98$, $t=2.09$, $pMCMC=0.020$) talkers. The model variance across passages approached zero, suggesting similar within-talker pause durations across passages. The relationship between articulation rate and pause duration is negative, demonstrating that talkers who spoke more slowly also produced longer pauses between phrases. The relationship between pause duration and number of pauses is also negative, confirming that talkers who paused more frequently produced shorter average pauses that talkers who paused less frequently. The main effect of gender reflects longer average pauses for male talkers (484 ms) than female talkers (417 ms). As shown in Table 1, the Southern talkers paused for longer on average than the talkers from the other dialect regions. These effects of gender and dialect are robust, even when articulation rate and pause frequency are included in the model, confirming the independent contribution of social factors to mean pause duration.

The speaking rate analysis revealed that the Southern talkers in the current study exhibited a slower overall articulation rate and longer overall pauses than the New England talkers. The Southern male talkers also exhibited the most frequent pauses of any dialect or gender group. These differences between the Southern and New England talkers are illustrated in Fig. 1. The Southern male talker produced the entire utterance with a slower articulation rate (5.03 syll/s) than the female New England talker (7.45 syll/s), and also paused between the two phrases, whereas the New England female talker did not. The Western and Midland talkers exhibited similar pause durations to the New England talkers, but the Midland talkers exhibited a similar speaking rate to the Southern talkers. Thus, although a clear distinction between the fast-talking New England talkers and the slow-talking Southern talkers emerged from this analysis, the other dialects combine features of both fast and slow speaking rate patterns found in the neighboring regions.

3.2. Consonant and vowel intervals

A summary of the measures of the consonant and vowel interval durations is shown in Table 2 for each dialect. For %*V*, the Southern dialect exhibited the largest proportion of vocalic intervals and the Western and Northern dialects exhibited the smallest proportion of vocalic intervals. The linear mixed effect model confirmed a significant difference between the Southern and Western dialects ($\beta=2.61$, $t=2.18$, $pMCMC=0.029$) and a marginal difference between the Southern and Northern dialects ($\beta=2.29$, $t=1.91$, $pMCMC=0.051$). The model variance across passages and talkers was similar in magnitude, suggesting that individual variation and textual variation contribute comparably to overall variation in %*V*.

For Δ*C*, the Midland dialect exhibited the greatest variability in consonant interval duration and the Southern dialect exhibited the least. The other four dialects exhibited comparable, intermediate degrees of variability. In addition to articulation rate, regional dialect, and talker gender, the linear mixed effects models for Δ*C* included %*V* as a fixed effect (see Frota & Vigario, 2001) and revealed significant effects of both %*V* ($\beta=-0.10$, $t=-17.77$, $pMCMC<0.001$) and articulation rate ($\beta=-5.71$, $t=-10.86$, $pMCMC<0.001$). However, none of the dialect differences were significant. The effect of gender and the interaction between dialect and gender were also not significant. The model variance across passages approached zero, suggesting similar within-talker variation in consonant interval duration across passages. For both %*V* and articulation rate, the relationship with Δ*C* was negative, suggesting less variable consonant intervals were produced when vowel intervals were relatively longer and when overall articulation rate was faster. The negative relationship between %*V* and Δ*C* is illustrated for the six regional dialects in Fig. 2. The negative relationship between articulation rate and Δ*C* is consistent with Dellwo and Wagner's (2003) observations that Δ*C* is rate-dependent. As articulation rate increases, the variability in consonant interval duration decreases.

For Δ*V*, the Southern and Midland dialects exhibited the greatest variability and the Northern and Western dialects exhibited the least variability. The linear mixed effects models revealed a significant effect of articulation rate ($\beta=-7.60$, $t=-11.92$, $pMCMC<0.001$), as well as independent significant differences between the Northern dialect and the Mid-Atlantic ($\beta=5.20$, $t=2.12$, $pMCMC=0.041$), Midland ($\beta=5.22$, $t=2.14$, $pMCMC=0.039$), and Southern ($\beta=7.85$, $t=3.24$, $pMCMC=0.002$) dialects and between the Southern and Western dialects ($\beta=7.71$, $t=3.20$, $pMCMC=0.002$). The effect of gender and the interaction between dialect and gender were not significant. The model variance across passages was small relative to the variance across talkers, suggesting similar within-talker variability in vowel interval duration across passages. As in the analysis of Δ*C*, the negative relationship between Δ*V* and articulation rate suggests that the duration of vowel intervals was less variable as articulation rate increased, consistent with the lack of normalization for speaking rate in this measure.

For *VarcoC*, the New England dialect exhibited the most variable consonant intervals and the Southern dialect exhibited the least variable consonant intervals. The linear mixed effects models confirmed this significant difference ($\beta=2.94$, $t=2.07$, $pMCMC=0.048$). The effect of gender and the interaction between dialect and gender were not significant. For *VarcoV*, the

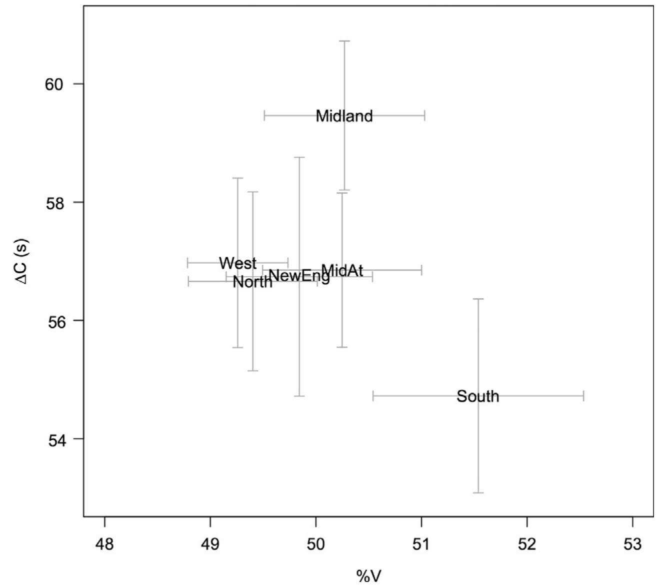

Fig. 2. Mean %V and ΔC for each dialect region. Error bars are standard error of subject means.

Southern talkers produced the most variable vowel intervals and the Northern talkers produced the least variable vowel intervals. The linear mixed effects models confirmed this significant difference between the Northern and Southern talkers ($\beta=3.89$, $t=2.19$, $pMCMC=0.037$). For both *VarcoC* and *VarcoV*, the model variance across passages was small relative to the variance across talkers, suggesting similar within-talker variability in consonant and vowel intervals across passages. Articulation rate was not a significant factor predicting *VarcoV*, suggesting that the normalization for mean interval duration was effective in reducing the impact of speaking rate on this measure of vowel interval variability. However, as reported by Dellwo (2006), articulation rate was a significant predictor of *VarcoC* ($\beta=1.61$, $t=3.46$, $pMCMC=0.001$), suggesting that the normalization for mean interval duration did not eliminate the impact of speaking rate on this measure of consonant interval variability.

For *rPVI-C*, the Midland and New England dialects exhibited the largest pairwise consonant interval variability and the Southern dialect exhibited the smallest. The linear mixed effects models confirmed a significant difference between the Southern and New England dialects ($\beta=5.20$, $t=2.13$, $pMCMC=0.039$), but the difference between the Southern and Midland dialects was not significant ($\beta=4.24$, $t=1.76$, $pMCMC=0.088$). Consistent with the lack of normalization for speaking rate in the calculation of *rPVI-C*, articulation rate was also a significant independent predictor of *rPVI-C* ($\beta=-8.93$, $t=-12.80$, $pMCMC<0.001$). As expected, the relationship between articulation rate and *rPVI-C* is negative, confirming that pairwise variability decreases as articulation rate increases. The effect of gender and the interaction between dialect and gender were not significant. The model variance across passages was small relative to the variance across talkers, suggesting similar within-talker variability in consonant intervals across passages.

For *nPVI-V*, the Western and Northern dialects exhibited the least pairwise vowel interval variability and the Southern dialect exhibited the most. The linear mixed effects models confirmed significant differences between the Southern dialect and the Northern ($\beta=3.52$, $t=2.14$, $pMCMC=0.035$) and Western ($\beta=4.14$, $t=2.54$, $pMCMC=0.017$) dialects. The effect of gender and the interaction between dialect and gender were not significant. In addition, articulation rate was not a significant factor predicting *nPVI-V*, suggesting that the normalization for mean interval duration was effective in reducing the impact of speaking rate on this measure of vowel interval variability. The variance across passages and talkers was similar in magnitude, suggesting that individual variation and textual variation contribute comparably to variation in pairwise vowel interval duration variation. Fig. 3 illustrates the relationship between *rPVI-C* and *nPVI-V* in these data.

Taken together, the consonant and vowel interval measures revealed less consonant variability for Southern talkers than for New England talkers, but more vowel variability for Southern talkers than for Northern and Western talkers. As with the speaking rate measures, the Southern and New England dialects emerge as unique in their temporal organization and the Midland and Mid-Atlantic dialects exhibit a more mixed profile.

4. Discussion

The results of our analysis extend previous research on speaking rate variability in regional dialects of American English through the combined analysis of articulation rate and pausing in a large number of dialects. This study also contributes new findings on regional variation in temporal organization across these six varieties. A summary of the significant effects of regional dialect and gender on temporal organization is shown in Table 3. Our speaking rate results are consistent with the stereotype that Southerners talk slowly and Northerners talk quickly. We observed a faster overall articulation rate for New England talkers than Southern and Midland talkers. This result, together with Jacewicz and colleagues' (Jacewicz et al., 2009, 2010) findings on articulation rates in

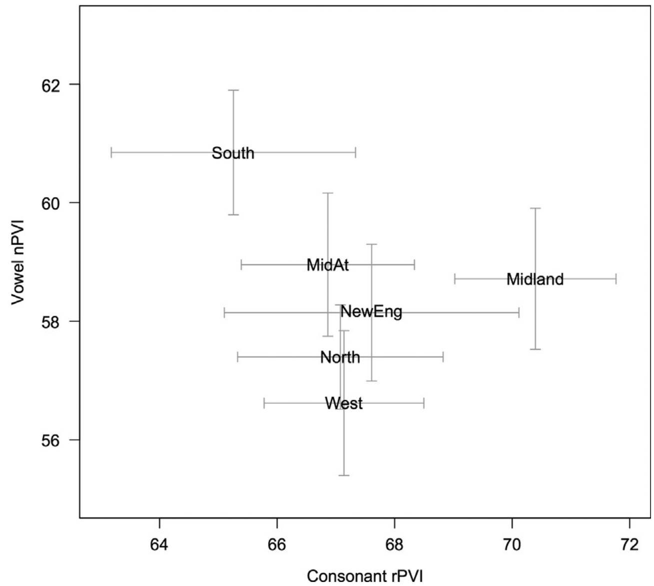

Fig. 3. Mean *rPVI-C* and *nPVI-V* for each dialect region. Error bars are standard error of subject means.

Table 3
Summary of the significant effects of regional dialect and gender on temporal organization.

Temporal property	Significant group differences
Articulation rate	New England > Midland, South
Pause frequency	Females: Midland > South
	Males: South > New England, Mid-Atlantic, North, Midland
Pause duration	South > New England, Midland, West
	Males > Females
ΔC	None
ΔV	Mid-Atlantic, Midland, South > North
	South > West
VarcoC	New England > South
VarcoV	South > North
rPVI-C	New England > South
nPVI-V	South > North, West

Wisconsin and North Carolina, suggest that the Northern stereotype may be based on speaking rates in both the northeastern and northern midwestern United States and that the Southern stereotype may also extend to talkers from the Midland dialect in the southern midwestern United States. The Southerners in our study also produced longer pauses than the New England, Midland, and Western talkers and the Southern male talkers paused more frequently than the male talkers from any other dialect region. The Midland female talkers also produced a large number of pauses, although the pauses themselves were not long, suggesting that the similarities between the temporal properties of Southern and Midland varieties extend beyond articulation rate. These results demonstrate that Southern speech is characterized not only by a slower articulation rate but also by longer and more frequent pauses, whereas New England speech is characterized by a faster articulation rate and shorter pauses (see also Byrd, 1994; Clopper & Smiljanic, 2011; Jacewicz et al., 2010; Kendall, 2009). Goldman-Eisler (1968) argued that the perception of speaking rate is a function of both articulation rate and the relative frequency and duration of pauses. Thus, the perception of slow-talking Southerners and fast-talking Northerners may reflect not only dialect differences in articulation rate between pauses, but also the temporal properties of the pauses themselves and the resulting prosodic phrasing.

The analyses of vowel and consonant interval duration variability, independent of speaking rate, provided further evidence of regional variation. The results revealed that Southerners have more variable vowel intervals than talkers from other regions and that Western and Northern talkers have more variable consonant intervals than talkers from other regions. Given Clopper et al.'s (2005) findings that the typically short lax vowels are relatively long in Southern speech, but that the typically long tense vowels are not longer in the South than in other regions, we might have expected to observe less vowel duration variability in the South than in the other dialects. Specifically, the longer lax vowels combined with typical tense vowels in the South should result in more similar tense and lax vowel durations overall. However, Clopper et al.'s (2005) results were based on the production of isolated hVd words and the current analysis was conducted with connected, read speech materials. Thus, the finding that Southern talkers produce more variable vowel intervals suggests that properties of running speech, such as vowel reduction, vowel–vowel hiatus resolution, and phrase-final lengthening, may vary substantially across dialects of American English. This interpretation is further supported by Byrd's (1994) observation of different vowel reduction patterns for Southerners than for Mid-Atlantic, Northern, and Western talkers. Clopper et al. (2005) also did not examine diphthongs, which are known to vary dramatically across dialects (Thomas, 2001). In addition, the lengthening and breaking of diphthongs and other long vowels in the South, which would be captured by the vowel interval variability measures in this study, may contribute to the perception of a Southern "drawl" (Thomas, 2001).

Crucially, the high degree of vowel variability observed in the Southern dialect cannot be attributed to its overall slower articulation rate. That is, the greater vowel duration variability is not simply the result of overall lengthening. All three of the vowel variability measures (ΔV, *VarcoV*, and *nPVI-V*) revealed significant differences between the Southern and Northern and/or Western talkers that were independent of the articulation rate covariate (ΔV) or observed despite inherent normalization for speaking rate in the dependent measure itself (*VarcoV*, *nPVI-V*). This independence between rate and vowel duration variability suggests that the slow-talking Southern stereotype may also be independent of the Southern drawl stereotype.

Similarly, the more variable consonant intervals for New England talkers may reflect differences across dialects in consonant cluster simplification, the realization of stop releases, flapping, and other consonant reduction phenomena. Although regional consonant variation in American English has not been as widely documented as vowel variation, Byrd (1994) reported less flapping among New England and Northern talkers than among Midland talkers. Given that flaps are typically short in duration, we would expect the realization of /t, d/ as stops in flapping environments to lead to less overall consonant variability in the Northern and New England dialects than the Midland dialect, in which the shorter flaps are produced, rather than the relatively high degree of consonant variability in the New England dialect that was actually observed. Flaps themselves are highly variable (Warner & Tucker, 2011), however, suggesting that a more detailed analysis of flapping across dialects may be necessary to understand its contribution to regional differences in consonant duration variability. Similarly, Byrd (1994) observed more glottal stops in vowel-initial words and in place of word-final alveolar stops among Northern and Southern talkers than among Midland talkers. Like flaps, glottal stops are typically short and may therefore increase consonant duration variability. The combined effects of flapping and glottal stops may therefore cancel each other out for the Northern and Midland talkers, but the increased glottal stopping among Southerners should lead to higher consonant variability, rather than the relatively low variability that was observed. Variation has also been observed across regional dialects in the realization of the coda consonant voicing contrast, including the relative duration of the preceding vowel and the coda consonant (Purnell, Salmons, & Tepeli, 2005; Royer, 2014).

A cursory inspection of our data, which were not collected with a large-scale analysis of consonant realization in mind, further reveals possible regional variation in the realization of coda stops and clusters. For example, the coda /t/s in the phrase *white light* were often deleted or realized by vowel glottalization in the New England dialect (11/20 tokens), whereas they were realized in the Northern dialect predominantly as unreleased stops (12/20). Similarly, the fricative-stop coda clusters in *lost* and *pushed* were primarily realized as a singleton fricative in the New England dialect (15/20 tokens), but were realized primarily as a fricative-stop sequence in the Western and Midland dialects (13/20 tokens each). The observed patterns of vowel and consonant duration variability across dialects in this study therefore point to the necessity of future research focusing on temporal aspects of regional vowel variation and on the documentation of regional consonant variation in American English, two areas that have been largely neglected in sociophonetic research (cf., Fox & Jacewicz, 2009).

More generally, the current findings demonstrate that the discussion of regional variation in American English must extend beyond vowel quality to include properties of temporal organization. The New England and Southern dialects emerged in this study as having the most distinctive temporal organization patterns, whereas the patterns for the Mid-Atlantic, Northern, Midland, and Western dialects were more mixed. These results contrast with descriptions of vowel variability in which the Northern, Southern, and Mid-Atlantic dialects are the most distinctive and the New England, Midland, and Western varieties have fewer characteristic properties and are therefore sometimes grouped together into a single "General American" dialect (Labov, 1998; Labov et al., 2006). Thus, although the Southern dialect has both distinctive vowels and distinctive temporal properties, the New England and Western dialects may be characterized more by temporal organization than by vowel quality and the Mid-Atlantic and Northern dialects may be characterized more by vowel quality than by temporal organization, at least for young, college-educated speakers of these dialects.

This orthogonal variation in vowel systems and temporal organization has theoretical implications for our understanding of American English regional dialects. First, the variability in temporal organization for the "General American" dialects that exhibit few characteristic vowel properties (i.e., New England, Midland, and Western) is striking. The New England talkers speak the fastest, but the Midland talkers speak slowly and are similar to the Southern talkers in terms of articulation rate. With respect to pause duration, however, the Midland and Western talkers are similar to the New England talkers and produce relatively short pauses overall. The New England dialect also exhibits markedly high variability in consonant duration, whereas the Western dialect exhibits markedly low variability in vowel duration. Thus, despite the similarities between these three dialects in terms of vowel quality, they exhibit significant differences in temporal organization, which suggests that their shared characterization as "General American" varieties may not be appropriate.

Second, the Western dialect shares low vowel duration variability with the Northern dialect, whereas the Midland dialect shares a slow speaking rate with the Southern dialect. The similarity in speaking rate in the Midland and Southern dialects is perhaps not surprising, given their geographic proximity and other shared features, such as back vowel fronting (Labov et al., 2006). However, the similarity in vowel duration variability between the Western and Northern dialects is unexpected. The vowel systems of these two varieties are very different (Clopper et al., 2005; Labov et al., 2006; Wright & Souza, 2012) and neither geography nor historical settlement patterns would lead to a prediction of shared temporal properties. The Western dialect is less well-documented overall than other regional varieties of American English and our results suggest that further research is essential for determining its relationship to the other better-described regional varieties of American English.

The results of the current study also have implications for our understanding of temporal properties such as speaking rate, pausing, and segmental duration variability, as well as the metrics that have been proposed to assess them. Articulation rate was included as a covariate in our analysis, so that all of the significant dialect differences observed in this study can be interpreted independently of overall speaking rate differences. Thus, although articulation rate was significantly correlated with pause frequency, pause duration, ΔC, ΔV, *VarcoC*, and *rPVI-C* dialect differences emerged for each of these measures except ΔC. These significant correlations with articulation rate are consistent with previous research (Dellwo, 2006; Dellwo & Wagner, 2003; Jacewicz et al., 2010). People who talk faster pause less frequently and for less time, resulting in an overall faster pace than people who talk slower, pause more, and for longer. In addition, faster speaking rates correspond to less variability in consonant and vowel intervals because there are limits on how short the shortest intervals can be (i.e., deletion). Thus, as speaking rate increases, the longest intervals are shortened more than the shortest intervals and the overall variability in duration decreases. Articulation rate was not correlated with the *VarcoV* or *nPVI-V* measures, suggesting that they effectively normalized for articulation rate (cf., Nokes & Hay, 2012; Torgersen & Szakay, 2012; Wiget et al., 2010). Note, however, that the ΔV results with overall articulation rate controlled statistically as a covariate were not identical to the *VarcoV* and *nPVI-V* results, suggesting that they capture slightly different aspects of the effects of speaking rate on vowel duration variability (i.e., overall articulation rate between pauses in the ΔV analyses, but relative duration of the consonant and vowel intervals in the *VarcoV* and PVI analyses). Similarly, although articulation rate was a significant covariate in all three of the consonant duration variability measures, the results for regional dialect differed slightly across the analyses. Thus, these measures not only reflect different aspects of the effect of speaking rate (as in the vowel analysis), but also potentially different aspects of consonant duration variability. Whereas *VarcoC* and *rPVI-C* revealed a significant difference between the New England and Southern dialects, ΔC did not. These results therefore provide additional evidence that these metrics are not simple variants of each other, but may actually reveal different aspects of the temporal organization of languages and dialects (Arvaniti, 2009, 2012; Loukina et al., 2011; Wiget et al. 2010). These results also dovetail on White et al.'s (2012) perceptual examination of within- and across-language variability, which showed that listeners can distinguish between varieties of British English on the basis of timing information alone. The current results, therefore, strongly predict that perceptual differentiation of American English varieties is at least partially dependent on consonant and vowel duration variation.

Unlike comparisons across languages and speaking styles (e.g., Arvaniti, 2009, 2012; Wiget et al., 2010), however, the results of our analysis comparing regional dialects within one language using a fixed set of speech materials produced highly consistent and interpretable results across the Varco and PVI metrics. Our data overall are also consistent with previous research on the temporal properties of English, with the possible exception of %*V*. Whereas in previous studies typically only 40% of the speech contained vowels (e.g., Ramus et al., 1999; Smiljanic & Bradlow, 2008; Wiget et al., 2010), our data contained 50% vowel intervals by duration. This larger %*V* in our study is more comparable to Arvaniti's (2012) results in which she observed 46% vowel intervals in English. The results of the other metrics also correspond well with the results reported for English by Arvaniti (2012), which may reflect more robust estimates of these temporal properties in Arvaniti's (2012) and our work due to the relatively large corpora (approximately 10 min per language/dialect) that were examined in these two studies.

The speech that we examined in our study consisted of two read passages produced by each talker. The passages differed in both overall length and in genre. The Rainbow passage is a shorter, pseudo-scientific text, whereas the Goldilocks passage is a longer fairytale. To explore the effect of passage on temporal organization, we included a random intercept for passage in all of our analyses and considered the variance in these intercepts across the two passages in each model. The random talker variance exceeded the random passage variance for all dependent variables except for pause frequency. We attributed this high passage variance in pause frequency to overall passage length. For most of the other dependent measures, the random talker variance exceeded the random passage variance by at least one order of magnitude. However, the random talker and passage variances were similar in magnitude for %*V* and *nPVI-V*. These results are consistent with Loukina, Rosner, Kochanski, Keane, and Shih's (2013) findings that talkers contribute more variability than texts to these temporal measures and that vowels contribute more than consonants to textual variability. These similar patterns of talker and textual variability across the two studies are particularly compelling given the different methods of annotation that were used. Whereas we segmented our data into consonant and vowel intervals by hand and assigned only coda approximants to vowel intervals, Loukina et al. (2013) relied on automatic segmentation processes and assigned all sonorant segments to vowel intervals. Taken together, the results of these two studies provide converging evidence for the roles of both group-level (language or dialect) and individual-level (talker) variation in temporal organization. However, they also suggest that effects of textual variation may be limited to particular measures of vowel variation, assuming that a sufficient quantity of speech is examined (see also Arvaniti, 2012; Wiget et al., 2010).

5. Conclusions

The Southern dialect is characterized by a slow overall speaking rate, long pauses, and highly variable vowel durations. The New England dialect is characterized by a fast overall speaking rate, short pauses, and relatively high variability in consonant durations. This distinction between slow-talking Southerners and fast-talking Northeasterners is consistent with social stereotypes about regional speaking rates in American English. The Midland, Northern, and Western dialects exhibit less distinctive patterns overall. Although the New England, Midland, and Western dialects are often characterized as “General American” varieties, the Midland dialect aligns with the Southern dialect in speaking rate and the Western dialect aligns with the Northern dialect in vowel duration variability. These findings demonstrate the important potential contribution of temporal properties in connected speech to the characterization of regional dialects in American English and to perceptual stereotypes relating to speaking rate and timing.

The analysis also demonstrated that the so-called “rhythm” metrics can be used to explore differences in timing across dialects within a language. Although regional variation in American English is not expected to reveal differences along the purported continuum from syllable- to stress-timing, the coefficient of variance and pairwise variability metrics revealed consistent differences in vowel timing patterns between the Southern and the Northern and Western dialects and in consonant timing patterns between the New England and the Southern dialects. These differences in timing patterns may have implications for the perception of variation, including stereotypes about phenomena such as the Southern “drawl”, and also suggest promising directions for future research, such as regional differences in consonant and vowel reduction processes. The current description of temporal variation in regional varieties of American English will serve as an essential backdrop for future explorations of the role of variation in temporal properties in the perception and representation of these dialects.

Acknowledgments

This work was supported by a Coca-Cola Critical Difference for Women Faculty Grant awarded to the first author. We would like to thank Victoria Cook, Christina King Garcia, and John Pate for assistance with data coding.

Appendix

See Tables 4–13.

Table 4
ArticulationRate~Dialect + IPI + (1|Passage) + (1|Talker).

Random effects:	Variance	SD						
Talker (Intercept)	0.163	0.404						
Passage (Intercept)	<0.001	<0.001						
	Dialect reference: Midland				Dialect reference: New England			
Fixed effects:	Estimate	SE	*t*	*pMCMC*	Estimate	SE	*t*	*pMCMC*
(Intercept)	5.279	0.139	38.00	0.007	5.653	0.139	40.75	0.021
MidAt	0.238	0.192	1.24	0.182	−0.136	0.193	−0.71	0.426
Midland					−0.374	0.192	−1.94	0.027
NewEng	0.374	0.192	1.94	0.033				
North	0.208	0.192	1.08	0.229	−0.166	0.192	−0.86	0.358
South	0.032	0.192	0.17	0.862	−0.342	0.192	−1.78	0.052
West	0.110	0.192	0.57	0.523	−0.264	0.192	−1.37	0.138
IPI	0.010	0.003	2.79	0.021	0.010	0.003	2.79	0.020

Table 5
NumberofPauses~Gender*Dialect + SyllSec + PauseDur + (1|Passage) + (1|Talker).

Random effects:	Variance	SD						
Talker (Intercept)	2.940	1.715						
Passage (Intercept)	16.682	4.084						
	Dialect reference: Mid-Atlantic				Dialect reference: South			
Fixed effects:	Estimate	SE	*t*	*pMCMC*	Estimate	SE	*t*	*pMCMC*
(Intercept)	30.055	5.600	5.37	<0.001	28.939	5.615	5.15	<0.001
Male	1.356	1.440	0.94	0.267	5.998	1.439	4.17	<0.001
MidAt					1.116	1.414	0.79	0.397
Midland	1.983	1.423	1.39	0.146	3.010	1.429	2.17	0.028
NewEng	0.224	1.435	0.16	0.864	1.341	1.457	0.92	0.335
North	0.892	1.412	0.63	0.494	2.008	1.413	1.42	0.130
South	−1.116	1.414	−0.79	0.400				
West	0.160	1.423	0.11	0.958	1.276	1.433	0.89	0.372
SyllSec	−2.640	0.685	−3.85	<0.001	−2.640	0.685	−3.85	<0.001
PauseDur	−8.207	4.402	−1.87	0.033	−8.207	4.402	−1.87	0.035
Male: MidAt					−4.642	2.003	−2.32	0.011
Male: Midland	−3.737	1.995	−1.87	0.043	−8.379	1.999	−4.19	<0.001
Male: NewEng	0.270	2.018	0.13	0.920	−4.372	1.998	−2.19	0.020
Male: North	−0.824	1.994	−0.41	0.646	−5.465	2.005	−2.73	0.004
Male: South	4.642	2.003	2.32	0.016				
Male: West	1.183	1.999	0.59	0.493	−3.459	2.012	−1.72	0.074

Table 6
PauseDuration~Gender+Dialect+SyllSec+NumPause+(1|Passage)+(1|Talker).

Random effects:	Variance	SD						
Talker (Intercept)	1251.10	35.37						
Passage (Intercept)	<0.01	<0.01						
	Dialect reference: New England				Dialect reference: South			
Fixed effects:	Estimate	SE	*t*	*pMCMC*	Estimate	SE	*t*	*pMCMC*
(Intercept)	696.75	80.63	8.64	<0.001	747.95	76.54	9.77	0.014
Male	70.13	11.90	5.89	<0.001	70.13	11.90	5.89	<0.001
MidAt	28.18	20.61	1.37	0.134	−23.02	20.78	−1.11	0.216
Midland	−1.33	21.12	−0.06	0.872	−52.53	20.56	−2.56	0.004
NewEng					−51.20	21.02	−2.44	0.009
North	31.29	20.71	1.51	0.094	−19.91	20.64	−0.97	0.276
South	51.20	21.02	2.44	0.007				
West	8.22	20.77	0.40	0.677	−42.98	20.60	−2.09	0.020
SyllSec	−48.98	13.43	−3.65	<0.001	−48.98	13.43	−3.65	<0.001
NumPause	−2.35	0.81	−2.92	0.048	−2.35	0.81	−2.92	0.045

Table 7
PercentV~Dialect+(1|Passage)+(1|Talker).

Random effects:	Variance	SD						
Talker (Intercept)	5.250	2.291						
Passage (Intercept)	1.595	1.263						
	Dialect reference: West				Dialect reference: South			
Fixed effects:	Estimate	SE	*t*	*pMCMC*	Estimate	SE	*t*	*pMCMC*
(Intercept)	49.045	1.232	39.83	0.003	51.655	1.232	41.94	0.024
MidAt	1.174	1.207	0.97	0.329	−1.436	1.204	−1.19	0.228
Midland	1.304	1.203	1.08	0.269	−1.307	1.200	−1.09	0.269
NewEng	0.596	1.206	0.49	0.611	−2.014	1.204	−1.67	0.092
North	0.317	1.205	0.26	0.785	−2.293	1.202	−1.91	0.051
South	2.610	1.199	2.18	0.032				
West					−2.610	1.199	−2.18	0.029

Table 8
DeltaC~Dialect+PercentV+SyllSec+(1|Passage)+(1|Talker).

Random effects:	Variance	SD						
Talker (Intercept)	5.607	2.368						
Passage (Intercept)	<0.001	<0.001						
	Dialect reference: South				Dialect reference: Midland			
Fixed effects:	Estimate	SE	*t*	*pMCMC*	Estimate	SE	*t*	*pMCMC*
(Intercept)	137.039	3.921	34.95	0.002	140.114	3.876	36.15	0.044
MidAt	1.340	1.814	0.74	0.460	−1.735	1.827	−0.95	0.353
Midland	3.075	1.795	1.71	0.093				
NewEng	2.259	1.817	1.24	0.222	−0.816	1.829	−0.45	0.647
North	0.889	1.805	0.49	0.624	−2.186	1.816	−1.20	0.235
South					−3.075	1.795	−1.71	0.099
West	−0.153	1.795	−0.08	0.932	−3.227	1.806	−1.79	0.078
PercentV	−0.996	0.056	−17.77	<0.001	−0.996	0.056	−17.77	<0.001
SyllSec	−5.706	0.525	−10.86	<0.001	−5.706	0.525	−10.86	<0.001

Table 9
DeltaV~Dialect+SyllSec+(1|Passage)+(1|Talker).

Random effects:	Variance	SD						
Talker (Intercept)	13.737	3.706						
Passage (Intercept)	4.397	2.097						
	Dialect reference: North				Dialect reference: South			
Fixed effects:	Estimate	SE	*t*	*pMCMC*	Estimate	SE	*t*	*pMCMC*
(Intercept)	95.323	4.198	22.71	0.066	103.173	4.107	25.12	0.034
MidAt	5.195	2.453	2.12	0.041	−2.655	2.435	−1.09	0.291
Midland	5.218	2.440	2.14	0.039	−2.633	2.416	−1.09	0.273
NewEng	3.524	2.451	1.44	0.158	−4.326	2.437	−1.78	0.084
North					−7.850	2.423	−3.24	0.002
South	7.850	2.423	3.24	0.002				
West	0.140	2.432	0.06	0.970	−7.710	2.411	−3.20	0.002
SyllSec	−7.595	0.637	−11.92	<0.001	−7.595	0.637	−11.92	<0.001

Table 10
VarcoC~Dialect+SyllSec+(1|Passage)+(1|Talker).

Random effects:	Variance	SD						
Talker (Intercept)	0.913	0.956						
Passage (Intercept)	<0.001	<0.001						
	Dialect reference: South				Dialect reference: New England			
Fixed effects:	Estimate	SE	*t*	*pMCMC*	Estimate	SE	*t*	*pMCMC*
(Intercept)	47.218	2.695	17.52	0.139	50.155	2.831	17.72	0.047
MidAt	1.249	1.421	0.88	0.397	−1.688	1.448	−1.17	0.259
Midland	2.319	1.400	1.66	0.103	−0.618	1.436	−0.43	0.690
NewEng	2.937	1.419	2.07	0.051				
North	2.771	1.405	1.97	0.054	−0.165	1.434	−0.12	0.896
South					−2.937	1.419	−2.07	0.044
West	1.467	1.394	1.05	0.304	−1.470	1.425	−1.03	0.320
SyllSec	1.613	0.466	3.46	0.001	1.613	0.466	3.46	<0.001

Table 11
VarcoV~Dialect+SyllSec+(1|Passage)+(1|Talker).

Random effects:	Variance	SD						
Talker (Intercept)	7.475	2.734						
Passage (Intercept)	<0.001	<0.001						
	Dialect reference: North				Dialect reference: South			
Fixed effects:	Estimate	SE	*t*	*pMCMC*	Estimate	SE	*t*	*pMCMC*
(Intercept)	56.363	2.874	19.61	0.047	60.254	2.803	21.50	0.027
MidAt	2.740	1.801	1.52	0.125	−1.151	1.788	−0.64	0.540
Midland	1.738	1.791	0.97	0.343	−2.153	1.774	−1.21	0.244
NewEng	2.019	1.799	1.12	0.255	−1.872	1.788	−1.05	0.325
North					−3.891	1.779	−2.19	0.040
South	3.891	1.779	2.19	0.033				
West	0.348	1.785	0.20	0.834	−3.543	1.770	−2.00	0.057
SyllSec	−0.486	0.466	−1.04	0.282	−0.486	0.466	−1.04	0.280

Table 12
rPVI-C~Dialect+SyllSec+(1|Passage)+(1|Talker).

Random effects:	Variance	SD						
Talker (Intercept)	10.199	3.194						
Passage (Intercept)	<0.001	<0.001						
	Dialect reference: South				Dialect reference: Midland			
Fixed effects:	Estimate	SE	*t*	*pMCMC*	Estimate	SE	*t*	*pMCMC*
(Intercept)	113.782	4.122	27.60	0.013	118.020	4.112	28.70	0.019
MidAt	2.244	2.440	0.92	0.368	−1.994	2.460	−0.81	0.423
Midland	4.239	2.416	1.76	0.090				
NewEng	5.201	2.440	2.13	0.039	0.963	2.461	0.39	0.707
North	3.249	2.423	1.34	0.191	−0.989	2.443	−0.41	0.679
South					−4.239	2.416	−1.76	0.085
West	2.346	2.409	0.97	0.344	−1.893	2.429	−0.78	0.443
SyllSec	−8.926	0.697	−12.80	<0.001	−8.926	0.697	−12.80	<0.001

Table 13
nPVI-V~Gender+Dialect+SyllSec+(1|Passage)+(1|Talker).

Random effects:	Variance	SD						
Talker (Intercept)	5.522	2.350						
Passage (Intercept)	3.290	1.814						
	Dialect reference: West				Dialect reference: South			
Fixed effects:	Estimate	SE	*t*	*pMCMC*	Estimate	SE	*t*	*pMCMC*
(Intercept)	58.044	3.028	19.17	0.142	62.181	3.020	20.59	0.035
Male	1.679	0.955	1.76	0.087	1.679	0.955	1.76	0.092
MidAt	2.812	1.657	1.70	0.097	−1.326	1.652	−0.80	0.436
Midland	1.958	1.645	1.19	0.249	−2.181	1.639	−1.33	0.189
NewEng	1.474	1.656	0.89	0.374	−2.664	1.652	−1.61	0.118
North	0.617	1.648	0.37	0.695	−3.522	1.642	−2.14	0.035
South	4.139	1.633	2.54	0.015				
West					−4.138	1.633	−2.54	0.018
SyllSec	−0.410	0.452	−0.91	0.376	−0.411	0.452	−0.91	0.374

References

Abercrombie, D. (1967). *Elements of acoustic phonetics*. Chicago: Aldine.
Arvaniti, A. (2009). Rhythm, timing, and the timing of rhythm. *Phonetica*, *66*, 48–63.
Arvaniti, A. (2012). The usefulness of metrics in the quantification of speech rhythm. *Journal of Phonetics*, *40*, 351–373.
Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language*, *59*, 390–412.
Byrd, D. (1994). Relations of sex and dialect to reduction. *Speech Communication*, *15*, 39–54.
Clopper, C. G., & Pisoni, D. B. (2006). The Nationwide Speech Project: A new corpus of American English dialects. *Speech Communication*, *48*, 633–644.
Clopper, C. G., Pisoni, D. B., & de Jong, K. (2005). Acoustic characteristics of the vowel systems of six regional varieties of American English. *Journal of the Acoustical Society of America*, *118*, 1661–1676.
Clopper, C. G., & Smiljanic, R. (2011). Effects of gender and regional dialect on prosodic patterns in American English. *Journal of Phonetics*, *39*, 237–245.
Coggshall, E. L. (2008). The prosodic rhythm of two varieties of Native American English. *University of Pennsylvania Working Papers in Linguistics*, *14*(2), 1–9.
Dauer, R. M. (1983). Stress-timing and syllable-timing reanalyzed. *Journal of Phonetics*, *11*, 51–62.
Dellwo, V. (2006). Rhythm and speech rate: A variation coefficient for deltaC. In P. Karnowski, & I. Szigeti (Eds.), *Language and language processing* (pp. 231–241). Frankfurt am Main: Peter Lang.
Dellwo, V., & Wagner, P. (2003). Relations between language rhythm and speech rate. In *Proceedings of the 15th international congress of phonetic sciences* (pp. 471–474).
Deterding, D. (2001). Letter to the editor: The measurement of rhythm: A comparison of Singapore and British English. *Journal of Phonetics*, *29*, 217–230.
Fagyal, Z. (2010). Rhythm types and the speech of working-class youth in a banlieue of Paris: The role of vowel elision and devoicing. In D. R. Preston, & N. Niedzielski (Eds.), *A reader in sociophonetics* (pp. 91–132). New York: De Gruyter Mouton.
Fairbanks, G. (1940). *Voice and articulation drillbook*. New York: Harper.
Fox, R. A., & Jacewicz, E. (2009). Cross-dialectal variation in formant dynamics of American English vowels. *Journal of the Acoustical Society of America*, *126*, 2603–2618.
Frota, S., & Vigario, M. (2001). On the correlates of rhythmic distinctions: The European/Brazilian Portuguese case. *Probus*, *13*, 247–275.
Goldman-Eisler, F. (1968). *Psycholinguistics: Experiments in spontaneous speech*. London: Academic Press.
Grabe, E., & Low, E. L. (2002). Acoustic correlates of rhythm class. In C. Gussenhoven, & N. Warner (Eds.), *Laboratory phonology*, 7 (pp. 515–546). Berlin: Mouton de Gruyter.
Henderson, A., Goldman-Eisler, F., & Skarbek, A. (1966). Sequential temporal patterns in spontaneous speech. *Language and Speech*, *9*, 207–216.
Holmes, J., & Ainsworth, H. (1996). Syllable-timing and Maori English. *Te Reo*, *39*, 75–84.
Holmes, J., & Ainsworth, H. (1997). Unpacking the research process: Investigating syllable-timing in New Zealand English. *Language Awareness*, *6*, 32–47.
Jacewicz, E., Fox, R. A., O'Neill, C., & Salmons, J. (2009). Articulation rate across dialect, age, and gender. *Language Variation and Change*, *21*, 233–256.
Jacewicz, E., Fox, R. A., & Wei, L. (2010). Between-speaker and within-speaker variation in speech tempo of American English. *Journal of the Acoustical Society of America*, *128*, 839–850.
Jacewicz, E., Salmons, J., & Fox, R. A. (2007). Vowel duration in three American English dialects. *American Speech*, *82*, 367–385.
Kendall, T. (2013). *Speech rate, pause, and sociolinguistic variation: Studies in corpus sociophonetics*. New York: Palgrave Macmillan.
Kendall, T. S. (2009). *Speech rate, pause, and linguistic variation: An examination through the sociolinguistic archive and analysis project* (Unpublished Ph.D. thesis). Durham, NC: Duke University.

Kowal, S., & O'Connell, D. C. (1980). Pausological research at Saint Louis University. In H. W. Dechert, & M. Raupach (Eds.), *Temporal variables in speech: Studies in honour of Frieda Goldman-Eisler* (pp. 61–66). The Hague: Mouton.
Labov, W. (1998). The three dialects of English. In M. D. Linn (Ed.), *Handbook of dialects and language variation* (pp. 39–81). San Diego: Academic Press.
Labov, W., Ash, S., & Boberg, C. (2006). *Atlas of North American English*. New York: Mouton de Gruyter.
Loukina, A., Kochanski, G., Rosner, B., Keane, E., & Shih, C. (2011). Rhythm measures and dimensions of durational variation in speech. *Journal of the Acoustical Society of America*, *129*, 3258–3270.
Loukina, A., Rosner, B., Kochanski, G., Keane, E., & Shih, C. (2013). What determines duration-based rhythm measures? Text or speaker?. *Laboratory Phonology*, *4*, 339–382.
Low, E. L., Grabe, E., & Nolan, F. (2000). Quantitative characterizations of speech rhythm: Syllable-timing in Singapore English. *Language and Speech*, *43*, 377–401.
Newman, M., & Wu, A. (2011). "Do you sound Asian when you speak English?" Racial identification and voice in Chinese and Korean Americans' English. *American Speech*, *86*, 152–178.
Niedzielski, N., & Preston, D. R. (2003). *Folk linguistics*. New York: Mouton de Gruyter.
Nokes, J., & Hay, J. (2012). Acoustic correlates of rhythm in New Zealand English: A diachronic study. *Language Variation and Change*, *24*, 1–31.
Nolan, F., & Asu, A. L. (2009). The pairwise variability index and coexisting rhythms in language. *Phonetica*, *66*, 64–77.
Purnell, T., Salmons, J., & Tepeli, D. (2005). German substrate effects in Wisconsin English: Evidence for final fortition. *American Speech*, *80*, 135–164.
Ramus, F., Dupoux, E., & Mehler, J. (2003). The psychological reality of rhythm classes: Perceptual studies. In *Proceedings of the 15th international congress of phonetic sciences* (pp. 337–342).
Ramus, F., Nespor, M., & Mehler, J. (1999). Correlates of linguistic rhythm in the speech signal. *Cognition*, *73*, 265–292.
Ray, G. B., & Zahn, C. J. (1990). Regional speech rates in the United States: A preliminary analysis. *Communication Research Reports*, *7*, 34–37.
Royer, A. J. (2014). *Effects of regional dialect on word-final consonant voicing* (Undergraduate thesis). Ohio State University.
Smiljanic, R., & Bradlow, A. R. (2008). Temporal organization of English clear and conversational speech. *Journal of the Acoustical Society of America*, *124*, 3171–3182.
Stockwell, P. (2002). *Sociolinguistics*. London: Routledge.
Szakay, A. (2006). Rhythm and pitch as markers of ethnicity in New Zealand English. In *Proceedings of the 11th Australian international conference on speech science and technology* (pp. 421–426).
Thomas, E. R. (2001). *An acoustic analysis of vowel variation in New World English*. Durham, NC: Duke University Press.
Thomas, E. R., & Carter, P. M. (2006). Prosodic rhythm and African American English. *English World-Wide*, *27*, 331–355.
Torgersen, E. N., & Szakay, A. (2012). An investigation of speech rhythm in London English. *Lingua*, *122*, 822–840.
Warner, N., & Tucker, B. V. (2011). Phonetic variability of stops and flaps in spontaneous and careful speech. *Journal of the Acoustical Society of America*, *130*, 1606–1617.
Warren, P. (1998). Timing patterns in New Zealand English rhythm. *Te Reo*, *41*, 80–93.
Wennerstrom, A., & Siegel, A. F. (2003). Keeping the floor in multiparty conversations: Intonation, syntax, and pause. *Discourse Processes*, *36*, 77–107.
White, L., & Mattys, S. L. (2007). Rhythmic typology and variation in first and second languages. In P. Prieto, J. Mascaro, & M.-J. Sole (Eds.), *Segmental and prosodic issues in Romance phonology* (pp. 237–257). Amsterdam: John Benjamins.
White, L., Mattys, S. L., & Wiget, L. (2012). Language categorization by adults is based on sensitivity to durational cues, not rhythm class. *Journal of Memory and Language*, *66*, 665–679.
Whiteside, S. P. (1996). Temporal-based acoustic-phonetic patterns in read speech: Some evidence for sex differences. *Journal of the International Phonetic Association*, *26*, 23–40.
Wiget, L., White, L., Schuppler, B., Grenon, I., Rauch, O., & Mattys, S. L. (2010). How stable are acoustic metrics of contrastive speech rhythm?. *Journal of the Acoustical Society of America*, *127*, 1559–1569.
Wright, R., & Souza, P. (2012). Comparing identification of standardized and regionally valid vowels. *Journal of Speech, Language, and Hearing Research*, *55*, 182–193.