Kent Academic Repository

Full text document (pdf)

Citation for published version

Smith, Rachel and Rathcke, Tamara V (2016) Glasgow Gloom or Leeds Glue? Dialect-Specific Vowel Duration Constrains Lexical Segmentation and Access. Phonetica (74). pp. 1-24. ISSN 1423-0321.

DOI

https://doi.org/10.1159/000444857

Link to record in KAR

http://kar.kent.ac.uk/56746/

Document Version

Author's Accepted Manuscript

Copyright & reuse

Content in the Kent Academic Repository is made available for research purposes. Unless otherwise stated all content is protected by copyright and in the absence of an open licence (eg Creative Commons), permissions for further reuse of content should be sought from the publisher, author or other copyright holder.

Versions of research

The version in the Kent Academic Repository may differ from the final published version.
Users are advised to check **http://kar.kent.ac.uk** for the status of the paper. **Users should always cite the published version of record.**

Enquiries

For any further enquiries regarding the licence status of this document, please contact:
researchsupport@kent.ac.uk

If you believe this document infringes copyright then please contact the KAR admin team with the take-down information provided at **http://kar.kent.ac.uk/contact.html**

Glasgow *gloom* or Leeds *glue*? Dialect-specific vowel duration constrains lexical segmentation and access

Rachel Smith[1*] and Tamara Rathcke[2]

[1]Glasgow University Laboratory of Phonetics, University of Glasgow, United Kingdom

[2]English Language and Linguistics, University of Kent, United Kingdom

* Corresponding author:

Dr. Rachel Smith

Glasgow University Laboratory of Phonetics

School of Critical Studies

University of Glasgow

12 University Gardens

Glasgow G12 8QQ

United Kingdom

Rachel.Smith@glasgow.ac.uk

Phone: +44 141 330 5533

http://www.gla.ac.uk/departments/englishlanguage/staff/rachelsmith/

Short title: Dialect-specific vowel duration and lexical access

Abstract

Timing cues are important in many aspects of speech processing, from identifying segments to locating word and phrase boundaries. They vary across accents, yet representation and processing of this variation is poorly-understood. We investigated whether an accent difference in vowel duration affects lexical segmentation and access.

In Glasgow English (GE), /i u e o/ are shorter than in Leeds English (LE), especially for /i u/ before voiced stops and nasals. In a word-spotting experiment, GE and LE participants heard nonsense sequences (e.g. *pobegloomezh*) containing embedded words (*gloom, glue*), with segmental qualities intermediate between GE and LE. Critical vowel durations were manipulated according to accent (*GE-appropriate* vowels shorter than *LE-appropriate* ones), and phonological context (vowels shortest before voiceless stops < voiced stops/nasals < voiced fricatives).

GE participants generally spotted words like *gloom* more accurately with GE-appropriate than LE-appropriate vowels. LE participants were less accurate than GE participants to spot words like *gloom* with GE-appropriate vowels, but more likely to spot embeddings like *glue*.

These results were broadly as predicted based on the accent differences, but depended less than expected on the accent-specific phonological constraints. We discuss theoretical implications regarding the representation of duration and the time course of lexical access.

1. Introduction

Speech in a regional accent other than one's own can be difficult to understand (Labov & Ash, 1997; Nathan, Wells & Donlan, 1998; Nathan & Wells, 2001; Floccia, Goslin, Girard, & Konopczynski, 2006; Clopper & Bradlow, 2008; Floccia, Butler, Goslin, & Ellis, 2009; Adank, Evans, Stuart-Smith, & Scott, 2009; Sumner & Samuel, 2009). This difficulty has many possible causes, including surprise at encountering a different variety of speech (Floccia et al., 2009) and attentional demands associated with encoding speaker identity (Magnuson & Nusbaum, 2007). One key factor among these may be the way the listener's perceptual phonetic category structure has developed as a function of their experience (e.g. Goldinger, 1996, 1998; Pierrehumbert, 2003, 2006). This study tests whether differences in contextually-conditioned vowel duration across regional accents of English affect the likelihood of words being successfully segmented and identified from continuous speech. If so, models of spoken word recognition must be adapted to accommodate experience-dependent variation in duration.

1.1. Perceptual sensitivity to spectral and durational aspects of accent variation

Early studies that identified cross-accent difficulties in speech comprehension tended to be general in focus, and did not seek to isolate particular phonetic differences that might be responsible for the difficulties (Floccia et al., 2006; Adank et al., 2009). More recently, converging evidence from a number of accents has emerged to support the view that word recognition is sensitive to accent-specific phonetic realisation and phonological organisation. That is, the same phonetic input may be processed differently by groups of listeners in systematic relation with the differences between their accents. For example, for rhotic speakers of North American accents, r-less variants of words like *baker* (e.g. [beɪkə]) do not prime *baker* as effectively as r-ful variants ([beɪkɚ]), while for non-rhotic speakers, both [beɪkɚ] and [beɪkə] prime *baker* (Sumner and Samuel, 2009). In French, where /e/ and /ɛ/ contrast word-finally in northern (including Standard French), but not in southern accents, Dufour, Nguyen, and Frauenfelder (2007) showed that southern French listeners treat words like /epe/ (*épée* , sword) and /epɛ/ (*épais,* thick) perceptually as homophones in a primed lexical decision task. Brunellière, Dufour and Nguyen (2011) followed up this finding with electrophysiological experiments using the mismatch negativity paradigm, showing that when hearing *épée* and *épais* northern French listeners have distinct amplitudes of electrical activity at scalp recording sites broadly consistent with processing of concrete and abstract words respectively, whereas southern French listeners do not show distinct patterns of amplitude as a function of the words' final vowels. Jacewicz and Fox (2012) demonstrated a range of vowel confusions when listener and speaker have different accents of (US) English, while Clopper, Pierrehumbert, and Tamati (2010) showed that cross-accent segmental confusability influences the processes of lexical competition among possible candidates for word recognition. Scott & Cutler (1984) demonstrated that segmental cues to phrase boundaries—flapping and palatalization—are interpreted differently by North American and British listeners in line with production patterns in the two accents.

Most of the studies reviewed above investigated spectral cues. In contrast, little is known about the perceptual consequences of accent differences involving durational cues. Durations reflect segmental identity, word boundaries, and phrasal boundaries, as well as rhythm, rate, and many other properties (see Fletcher, 2010, for an extensive review). Experiments show that listeners are exquisitely sensitive to durational variation, and integrate durational information over a surprisingly extensive range of timescales to guide their perceptual decisions concerning segmental identity (Ainsworth 1972; Mermelstein 1978; Strange 1989; Whalen 1989; Fischer & Ohde, 1990), word boundary location (Davis et al., 2002; Salverda et al., 2003; Cho, McQueen, & Cox, 2007; Dilley & McAuley, 2008), morphemic structure (Kemps, Ernestus, Schreuder & Baayen, 2005) and phrase boundary location (Lehiste, Olive, & Streeter, 1972; Scott, 1982; Christophe, Peperkamp, Pallier, Block & Mehler, 2004). To the extent that durational patterning varies across accents, therefore, listeners' experience with their native accent may affect the way durational information is interpreted during on-line speech perception.

Investigation of this issue has to date focused on the level of segmental identity. Miller, Mondini, Grosjean, and Dommergues (2011) investigated the durations of perceived 'best exemplars' for /o/ and /ɔ/ (as in *côte* and *cotte*) for listeners whose native accent is Swiss French (where there is a large durational difference between these vowels) compared to Standard French (where the durational difference is very small). Durational variation in the stimulus affected best exemplar perception for both accents (contra the earlier study of Miller

& Grosjean, 1997) but much more strongly for Swiss French. As far as we know there is no research that investigates how accent variation might affect listeners' ability to use context-sensitive durational information to make perceptual decisions beyond segmental identity, e.g. to guide their segmentation and recognition of words. Such research is of high theoretical interest given ongoing debate about both the way phonetic details and contextual dependencies are represented in memory (e.g., Goldinger, 1996, 1998; Norris et al., 2003; Johnson, 2006; Hawkins, 2003; Norris, Cutler, & McQueen, 2003) and about the role of time in lexical competition processes (e.g. Magnuson, Dixon, Tanenhaus, & Aslin, 2007). As outlined below, comparison of the durational characteristics of the vowel systems of varieties of English spoken in Glasgow (in lowland Scotland) and Leeds (in Northern England) indicates a rich and complex pattern of variability that may have significant consequences for word recognition.

1.2. Durational variation in Glasgow and Leeds vowels

The present study used Glasgow English (GE) and Leeds English (LE) as accents representative of large urban areas in Scotland and Northern England respectively. For present purposes, we defined GE and LE as Standard English spoken with the urban accents of each city. As neither is very prominent in the UK national broadcast media, each can be expected to be equivalently unfamiliar to listeners who speak the other accent, an important consideration since familiarity acquired via the broadcast media is thought to affect cross-accent speech comprehension (e.g. Floccia et al, 2006; Adank et al, 2009).

The focus of this study is the difference between the two accents in terms of vowel duration. The available evidence suggests that /i u e o/ (among other vowels) are phonetically shorter in GE than LE (Wells, 1982; McKenna, 1988; Agutter, 1988; Ladd et al. 2009). There are also well-established differences in the way vowel duration is conditioned by phonological context. Table 1 shows a simplified view of the main patterns.

In LE, as in most other English accents and many other languages, vowels are longer prior to a voiced compared with a voiceless obstruent. Thus /i/ is longer in *seed* than in *seat* and, as fricatives condition longer duration than stops, /i/ is slightly longer still in *seize* (House & Fairbanks, 1953; Keating, 1985; Peterson & Lehiste, 1960). It is also longer when in an open than a closed syllable, and when the vowel directly precedes a word or morpheme boundary (e.g. *see;* Beckman & Edwards, 1990; Berkovits, 1994; Wightman, Shattuck-Hufnagel, Ostendorf & Price, 1992).

In GE, vowel durations are affected differently by a following consonant, according to what is known as the *Scottish Vowel Length Rule* or *SVLR* (Aitken, 1981; Scobbie, Hewlett & Turk, 1999; Agutter, 1988). The SVLR applies to the close vowels /i/ and /u/ and to the diphthong /aɪ/ (Scobbie et al., 1999); we term these "strong-SVLR" vowels, in contrast to "weak-SVLR" vowels which may show evidence of the rule in vernacular Scots dialects, but do not do so in Standard Scottish English (Johnston, 1997; Aitken, 1981). For strong-SVLR vowels, voiced stops, nasals, and /l/ all condition relatively short duration, though not quite as short as voiceless consonants, i.e. vowels in these contexts are slightly longer than vowels before voiceless consonants (McKenna, 1988; Scobbie et al., 1999). Voiced fricatives /v ð z ʒ/ and /r/ are the only singleton syllable-final consonants that condition substantially longer duration than voiceless consonants do. Long duration is also conditioned by open syllables, word boundaries and morpheme boundaries, as in LE. A consequence in GE is what Scobbie and Stuart-Smith (2008) term quasi-phonemic contrasts, i.e. morphologically-conditioned distinctions in a few word pairs, e.g. *brood—brewed,* where the tautomorphemic /d/ of *brood* conditions short /u/, while the heteromorphemic /d/ of *brewed* conditions long /u/. LE, along with other English English accents, lacks these distinctions, because both tautomorphemic and heteromorphemic /d/ condition long /u/.

Following context	Example	LE pattern	GE pattern
Voiceless stops and fricatives	*seat*	short	short
Voiced stops, nasals and /l/	*seed*	long	short
Voiced fricatives and /r/	*seize*	long	long
Word or morpheme boundary	*see*	long	long

Table 1. Typical patterns of vowel duration according to following phonological context in Leeds English (LE) and Glasgow English (GE), adapted from Scobbie et al. (1999). The terms "short" and "long" suggest a binary distinction; in fact, for both accents, vowel duration increases successively from the top to the bottom row of the table, but the difference between successive "short" rows is small, whereas the difference between "short" and "long" rows is large. See text for additional nuance.

In summary, GE exhibits the SVLR, while LE exhibits the typical English pattern whereby all voiced consonants condition longer duration than voiceless consonants. Unlike more northerly English varieties spoken close to the Scottish border (such as Newcastle and Berwick-upon-Tweed: Milroy, 1995, Watt & Ingham 2000), there is no evidence of SVLR-like behaviour in Leeds currently or even historically (Joan Beal, personal communication).

GE and LE also differ in other aspects of their vowel systems. The GE system (Abercrombie, 1979; Stuart-Smith, 1999, 2003, 2004) has a smaller inventory, with nine monophthongs and three diphthongs compared to the 13 monophthongs and six diphthongs characteristic of LE (Wells, 1982: 364-5). Further, the phonetic qualities of many vowels in the respective systems differ, though the accents also share some commonalities, e.g. both allow monophthongal realisations of the vowels in FACE and GOAT (Wells, 1982) which are diphthongal in southern English accents. For the purposes of this study, the key points are that both GE and LE have phonemic categories /i/, /u/, /e/ and /o/ (corresponding to the vowels of FLEECE, GOOSE, FACE and GOAT respectively) and that the main differences in these vowels across the accents lie in quantity not quality. Typical phonetic realisations are reported to be [i ~ i̞], [ʉ̞], [e] and [o] for standard GE (Stuart-Smith, 1999) and [iː], [uː], [eː (~ ɛɪ)] and [oː ~ ö̞ː (~ ɔʊ)] for LE (Wells, 1982). Our own observations suggest a more open quality for LE /e/ and /o/ than Wells' transcriptions do, together with a somewhat fronted quality for /u/ in some contexts (cf. Ferragne & Pellegrino, 2010 on /u/-fronting elsewhere in Northern England).

1.3. Perceptual implications of Glasgow and Leeds vowel duration patterns

There is little research on how durational differences between English accents are perceived. The few existing studies on perception of the SVLR focus on the perception of quasi-phonemic contrasts conditioned by morphology. Ferragne and colleagues have shown that compared to French listeners, Scottish listeners have steeper identification functions in response to a durational continuum from *brood* to *brewed* (Ferragne, Bedoin, Boulenger & Pellegrino 2011) and also differences in the P3a component in an oddball paradigm (Boulenger, Ferragne, Bedoin & Pellegrino, 2011). They conclude that these differences reflect that Scottish listeners treat the *brood—brewed* contrast as phonemic.

Given that durational cues subserve many linguistic functions (see 1.1), we might expect duration to convey conflicting messages across accents. In particular, duration is important in signalling word boundary location (Davis et al., 2002; Salverda et al., 2003; Cho, McQueen, & Cox, 2007; Dilley & McAuley, 2008), along with a range of other factors such as language-specific rhythm (Cutler & Norris, 1988), phonotactics (McQueen, 1998), and allophonic detail (Smith & Hawkins, 2000). No research, however, has addressed how lexical segmentation or lexical access from continuous speech might be affected by cross-accent differences in vowel duration related to the SVLR. Such an investigation holds interest: not only may accent-specific vowel durations act as more or less effective cues to the relevant segmental identities, and consequently to representations of words containing those segments, but they may also act differently as cues to lexical segmentation, since vowel duration typically varies with position in syllable and word (Maddieson, 1985).

1.4. Aims and predictions of the current study

In this study, the word-spotting paradigm was used to investigate whether accent-inappropriate vowel duration disrupts lexical segmentation and access for speakers of GE and LE. Word-spotting has been extensively used to investigate word segmentation (e.g. Cutler & Norris, 1988; McQueen, 1998): listeners hear nonsense strings that have a real word embedded in them, and make a speeded button press if they spot a word, which they then have to produce.

The study was guided by the following research questions:

1a) Most generally, we asked whether accent-inappropriate vowel duration would disrupt the accuracy and/or speed of word-spotting. In contrast to studies which demonstrate that unfamiliar accents disrupt lexical access but do not probe the phonetic bases of the disruption (e.g. Floccia et al, 2006, 2009; Adank et al, 2009), we wished to isolate the contribution of accent-specific vowel duration from that of other phonetic properties relevant to word segmentation, such as spectral cues (Thiessen & Saffran, 2004) and intonation (Welby, 2009) which may also vary according to accent. To this end, base stimuli were produced by a trained phonetician to have segmental targets intermediate between those of GE and LE. The duration of critical vowels was then manipulated artificially so as to be appropriate for one or the other accent. We predicted disruption to occur for both LE and GE listeners when vowel duration was inappropriate for their accent, though perhaps more subtly than in Floccia and Adank's studies since only a single cue was manipulated here. Disruption was assessed in two ways: first by within-accent comparisons, i.e. whether listeners from each group performed worse on accent-inappropriate durations than accent-appropriate ones; and second by between-accent comparisons, i.e. whether GE listeners performed worse than LE listeners on LE-appropriate vowel durations, and vice versa.

1b) More specifically, we sought to investigate the possible dependence of word-spotting responses upon the details of accent-specific contextual conditioning of duration. We expected to see the largest divergence across the accents in contexts where the accents clearly obey different patterns of phonological conditioning, namely in the context of voiced stops and nasals, which condition relatively short duration in affected vowels of GE, but long duration in LE vowels. Accordingly, we manipulated phonological contexts to be *non-lengthening* (i.e. followed by a voiceless stop or fricative); *moderate-lengthening* (i.e. followed by voiced stops and nasals) and *maximal-lengthening* (i.e. followed by voiced fricatives). Additionally, vowel type was manipulated, contrasting strong-SVLR vowels (/i/ and /u/) with a control condition of weak-SVLR vowels (/e/ and /o/).

2) Finally, we also considered the effect of accent-specific vowel duration on competition between alternative lexical embeddings, such as *glue* and *gloom* in the sequence *pobegloomezh* (/pɒbəˈgluməʒ/). While it is uncommon in word-spotting experiments to embed more than one real word in a stimulus, doing so was inevitable for this study because English has prohibitively few words of CVC structure that contain the vowels /i u e o/ and lack an embedded CV word. Accordingly, when we presented listeners with words like *pobegloomezh,* we had to count *gloom* responses and *glue* responses separately (see section 3.1). We capitalised on this aspect of the English lexicon to generate our second set of predictions, as follows:

2a) In general, we would expect listeners to be more likely to spot the longest word that is consistent with the input, rather than a shorter embedded word, i.e. to spot *gloom* more often than *glue.* This "long-word bias" has been found in previous experiments and supported by computational modelling (Davis, Gaskell, & Marslen-Wilson, 2002, Frauenfelder & Peeters, 1990; Pitt & Samuel 2006).

2b) However, this prediction is modulated by vowel duration. Vowel duration tends to be greater syllable- and word-finally than –internally (Maddieson, 1985). Therefore, if a vowel sounds inappropriately long, listeners will be less likely to assume it is word-internal, and more likely to interpret it as evidence of an upcoming word boundary (i.e., to spot *glue* rather than *gloom*). In our design, LE-appropriate durations were always longer than GE-appropriate durations, so GE listeners were the only group who heard inappropriately long durations. Thus, we arrive at an asymmetric prediction: When vowel duration is LE-appropriate, we expected GE listeners to spot more vowel-final words, whereas LE listeners should continue to show the ordinary long-word bias in this situation.

2. Method

2.1. Participants

There were 78 participants: 39 speakers of GE (28 male, mean age 19.9 years), and 39 speakers of LE (19 male, mean age 20.9 years). All were monolingual, with normal speech and hearing. GE participants had grown up in Greater Glasgow and LE participants in Greater Leeds, without spending more than 3 years resident outside the respective accent area. No GE participants had resided in Yorkshire, nor LE participants in Scotland, and none had a parent from the other accent area. Participants received a small fee for their participation.

2.2. Materials

The 66 experimental stimuli (Appendix) were nonsense sequences, such as *flizoomip, bezifreakib,* each of which had a *target sequence* embedded within it. Each target sequence was a (C)CVC sequence that contained two real word targets. The onset-vowel-coda or *OVC target* corresponded to the entire target sequence (e.g. *zoom, freak*) while the onset-vowel or *OV target* (e.g. *zoo, free*) corresponded to the initial consonant(s) and vowel of the target sequence. Other lexical embeddings were avoided where possible, but in a few cases it was necessary to allow an extra real word that overlapped partially with the onset of the OVC target, e.g. *rue* in *kigroovip* (OVC target *groove,* OV target *grew*). These extra embedded words were expected to be spotted rarely because of the difficulty of breaking up an onset cluster (McQueen, 1998; Smith & Hawkins, 2000). The 66 target sequences fell into six groups (11 each) according to their vowel type (strong-SVLR, weak-SVLR) and phonological context (non-lengthening, moderate-lengthening, and maximal-lengthening). Differences in lexical frequency across the experimental set were unavoidable given the phonological constraints on the choice of materials, and were dealt with statistically.

Nonsense strings had 3 syllables (wSw, n=33) or 4 syllables (swSw, n=33; S = primary stress, s = secondary stress, w = weak syllable). Thus, the target sequence was always the penultimate syllable and bore primary stress. The use of strings with more than two nonsense syllables is rather rare in word-spotting (though cf. Kim and Cho, 2009) and was intended to give listeners a sufficient preceding temporal context for interpretation of vowel duration. The number of syllables preceding the target varied to prevent attention from being drawn only to the second syllable. The weak syllables abutting the target word always contained a reduced vowel, /ə/ or /ɪ/. In 4-syllable nonsense strings, the first (secondary-stressed) syllable contained a lax vowel (/ɪ ɛ a ɒ ʌ/).

There were 66 filler items, with the prosodic structure wSw (n=33) or swSw (n=33). 33 fillers had /i u e o/ in their primary stressed syllable, and the remaining 33 had a lax vowel. They contained no real word or (in 10 cases, for parity with a few experimental items with extra embeddings) a word embedded inside a consonant cluster, e.g. *woo* in /fapəˈtwunəg/).

2.3. Recordings

The OVC and OV target words were recorded in a carrier sentence by a male speaker of GE and a female speaker of LE; both were lower middle-class speakers in their 20s. To inform decisions about implementation of the durational manipulation, segmental durations were measured in Praat.

The production of the base stimuli in an intermediate accent proceeded as follows. A male phonetician with expertise in accents of English served as the speaker. His regional background was mixed, including periods of residence in both the mid-north and south-east of England, but not in Scotland. He listened to the LE and GE recordings to familiarise himself with the speakers' accents. Also available to the speaker were IPA representations of the main variant(s) for each vowel for each accent, based on narrow transcriptions made by the first author: these were [(ɪ)i̠̞], [ʉ̞ ~ ʉ], [ɛ̞], and [o̞̜] for /i/, /u/, /e/, and /o/ in LE, and [i], [ʏ ~ ɵ], [e̞], and [o] for /i/, /u/, /e/, and /o/ in GE. The speaker then read the experimental and filler nonsense strings aloud from a list. Seven tokens of each item were produced: one token in imitation of the LE accent, one in imitation of the GE accent, and five intermediate-accenttokens that aimed to achieve segmental qualities between the two accents. Producing the intermediate-accent variants was a difficult but doable task for the speaker. The first author listened to the recordings and asked the speaker to re-record any problematic items. For each item, the first author selected (from the five intermediate-accent tokens recorded) one fluently-spoken token that most successfully achieved the desired intermediate qualities, for use in the perception tests.

2.4. Evaluation of intermediate accent

We evaluated how successfully the speaker had produced intermediate qualities in two ways, via acoustic analysis of the target vowels and via an accent rating pre-test.

2.4.1. Acoustic analysis

We analysed the target words spoken by the phonetician in all 66 selected intermediate tokens, plus a subset of 10 imitated-LE and 10 imitated-GE tokens, and also all 66 tokens spoken in the carrier sentence by each of the LE and GE speakers (n = 218 tokens in total). F1 and F2 of the target words' vowels were measured at vowel midpoint using LPC analysis in Praat (Burg method, window length 25 ms, 5 expected formants, max formant 5000 Hz) together with hand-correction in Formant Editor v.0.8.2 (Soskúthy, 2015). The formant values were normalised using Lobanov's z-score method ($F_i^N = (F_i - \mu_i)/\sigma_i$; Lobanov, 1971). This method successfully deals with gender-based variation (the LE speaker was female and the others male), though it may also erode some aspects of accent-based variation (cf. Adank, Smits, & van Hout, 2004).

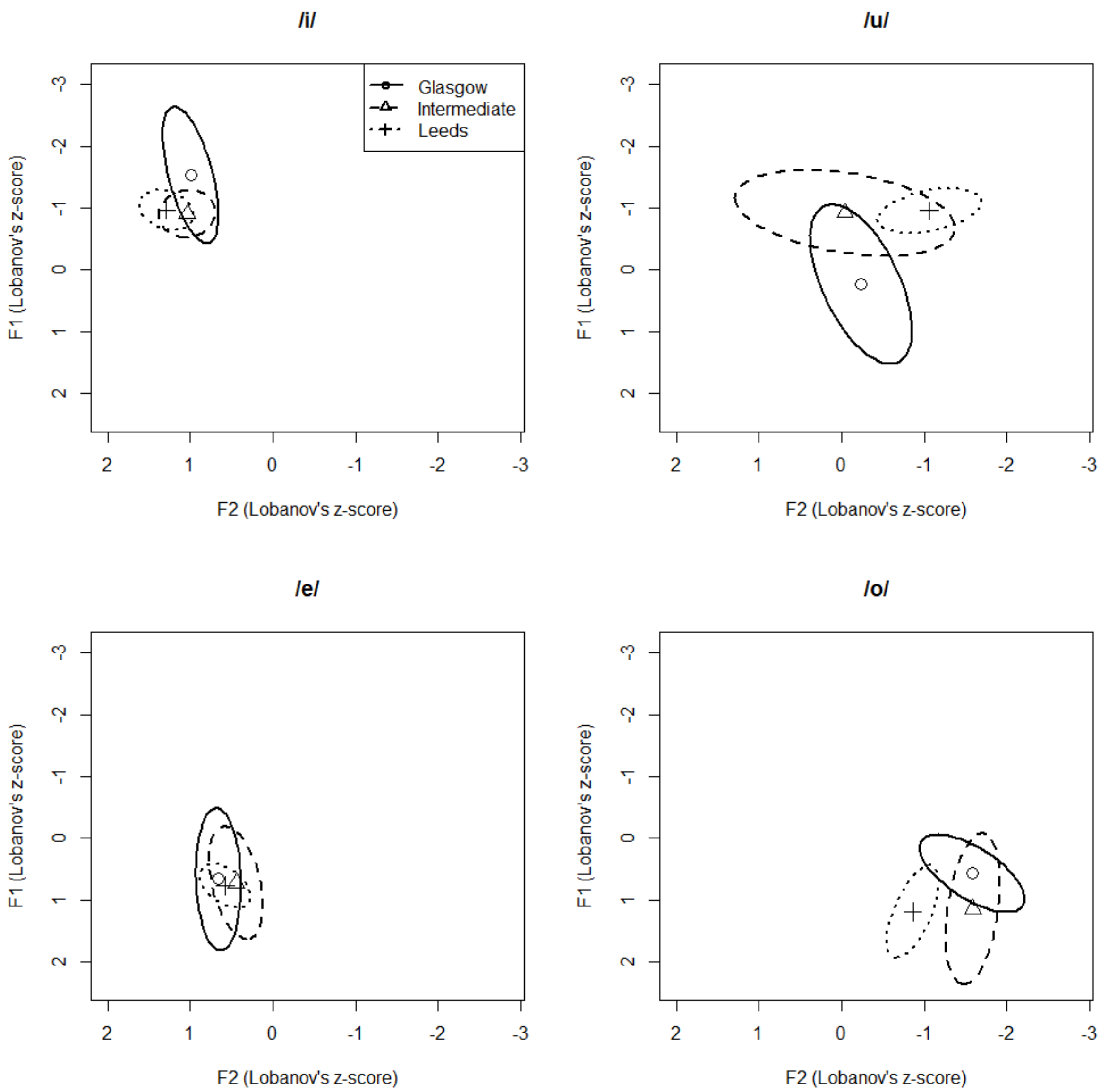

Figure 1. F1 and F2 of /i/, /u/, /e/ and /o/, transformed using Lobanov's z-score method (Lobanov, 1971). Points show the mean for each vowel category, and ellipses contain 90% of the data points, for the LE speaker, the GE speaker and the phonetician's Intermediate tokens.

Figure 1 shows the results. Comparison of the phonetician's intermediate tokens with the authentic GE and LE productions suggests success in achieving an intermediate accent. One vowel, /e/, appears very similar in all three accents (if somewhat less variable in LE); note however that the Lobanov transform has removed some accentual variation for /e/, which (together with /o/) is phonetically closer in GE than LE. For the other three vowels, /i/, /o/, and /u/, the intermediate tokens are, as expected, located between the GE and LE tokens in normalised F1-F2 space. Intermediate /i/ is slightly closer to LE than GE, and intermediate /o/ slightly closer to GE than LE, while for /u/, the intermediate value is close to LE on the height dimension (i.e. normalised F1), but close to GE on the frontness dimension (i.e. normalised F2). Within-category variability tends to be greater in the intermediate tokens than in the authentic accents, which probably reflects the difficulty of the task of producing them. Taking the four critical vowels together, the data clearly indicate that the speaker achieved a vowel space that is neither that of LE or GE, but in between.

2.4.2. Accent rating pre-test

The accent rating pre-test was conducted to evaluate the success of the intermediate accent from a perceptual point of view. Participants were 14 speakers of GE (10 male; mean age 25.2 years), tested in Glasgow, and 15 speakers of LE (8 male; mean age 23.2 years), tested in Leeds. None participated in the main experiment. Stimuli were the intermediate-accent tokens of all 66 experimental items, plus the 10 imitated-LE and 10 imitated-GE tokens that were used in the acoustic analysis (86 stimuli in total). A randomised list of the stimuli was presented using DMDX (Forster & Forster, 2003) with an inter-stimulus interval (ISI) of 2.5 s. Participants rated how similar the speaker's regional accent sounded to their own accent, on a scale from 1 (not at all close) to 7 (very close).

	Intermediate-accent	Imitated-GE	Imitated-LE
GE listeners	3.51 (1.53)	3.72 (1.77)	3.25 (1.54)
LE listeners	3.55 (1.76)	3.26 (1.90)	4.20 (2.09)

Table 2. GE and LE listeners' mean (sd in parentheses) ratings of intermediate-accent and imitated-accent stimuli, on a scale from 1 ("not at all close to your own accent") to 7 ("very close to your own accent").

Table 2 shows that both GE and LE listeners accorded the intermediate-accent stimuli very similar ratings in the middle of the 7-point scale (mean 3.51 for GE, 3.55 for LE). Thus, both listener groups judged these stimuli to be neither very close to, nor very far from, their own speech variety. Results for the imitated-GE and imitated-LE stimuli show that as expected, each listener group gave higher ("closer to your own accent") ratings to the speaker's imitations of their own than of the other accent. LE listeners rated imitated-LE stimuli as slightly closer to their own accent than GE listeners rated imitated-GE stimuli. Note however that this does not imply that the speaker's *intermediate* stimuli were necessarily any closer to LE than GE.

Statistical tests were carried out only on the responses to the intermediate-accent stimuli, since these alone were used in the main experiment. The data were submitted to linear mixed effects modelling (Baayen, 2008). First, a saturated model was fitted with fixed factors of Listener Accent (GE/LE), Context (Non-lengthening, Moderate-lengthening, and Maximal-lengthening), and Vowel (/i/, /u/, /e/, and /o/) and all their interactions, and with Subject and Word as random factors. Non-significant predictors were sequentially removed until the simplest model had been found. There was no significant main effect of Listener Accent, nor any significant interactions involving Listener Accent: that is, across all contexts and vowels, the GE and LE listeners did not differ in how close to their own accent they perceived the intermediate-accent stimuli to be. No other main effects or interactions were significant.

Based on the acoustic data and the rating pre-test, we conclude that the intermediate-accent stimuli are appropriate base stimuli to test the effect of accent-appropriate vowel duration in isolation from other accent characteristics.

2.5. Stimulus manipulations

PSOLA resynthesis in Praat (Boersma & Weenink, 2010) was used to create two versions (*GE-appropriate* and *LE-appropriate*) of each intermediate-accent stimulus, differing only in the duration of the vowel. The durations of *GE-appropriate* versions were set to be shorter than those for *LE-appropriate* versions, based on pilot production data from the two accents (see above). Figure 2 shows the implemented *GE-appropriate* and *LE-appropriate* durations. Note that the implemented values of *GE-appropriate* and *LE-appropriate* varied according to both vowel type and phonological context. In accordance with descriptions of the two varieties: *GE-appropriate* (short) durations are always shorter than *LE-appropriate* (long) durations for all vowels and contexts, but the largest *LE-minus-GE-appropriate* difference occurs for SVLR-vowels in moderate-lengthening contexts (e.g. *zoom, seed*) where short allophones are found in GE, and long ones in LE.

The pitch contour of the tokens was normalized using PSOLA to achieve a rise-fall on the nuclear stressed syllable of the target sequence. For the example *flizoomip*, f0 was set to 105 Hz at the start of the /flɪ/ syllable. It fell to 100 Hz at 10% of the way through /u/, then rose to 115 Hz at 90% of the way through this vowel, before falling to reach 90 Hz 30 ms after the start of /ɪ/ in the final syllable, and then to 85 Hz at the end of the final syllable. For four-syllable items (e.g. *bezifreakib*, the procedure was identical except that f0 was additionally set to 110 Hz at the start of the initial /bɛz/ syllable. The manipulation ensured that any intonational cues to accent were also removed.

2.6. Procedure

Two lists were prepared, each containing *LE-appropriate* versions of 33 experimental stimuli and *GE-appropriate* versions of the other 33 (counterbalanced for vowel type and context), plus all 66 fillers. There were four different randomised versions of each list. Participants were assigned randomly to lists and versions.

LE and GE participants were tested in soundproof booths at the Universities of Leeds and Glasgow respectively. DMDX was used to present the experiment and record button presses and spoken responses.

Participants held an Xbox® controller in their dominant hand. They were instructed that on each trial they would hear a sequence of nonsense syllables, sometimes with a real word within the sequence. No comments were made on the accent of the stimulus speaker. Upon hearing a real word, they should press the button with their index finger as fast as possible, then speak the word aloud. The inter-stimulus interval was 4 sec.

There was a practice session of ten items (without feedback), after which participants could ask questions, followed by the main test. The experiment lasted approximately ten minutes in total.

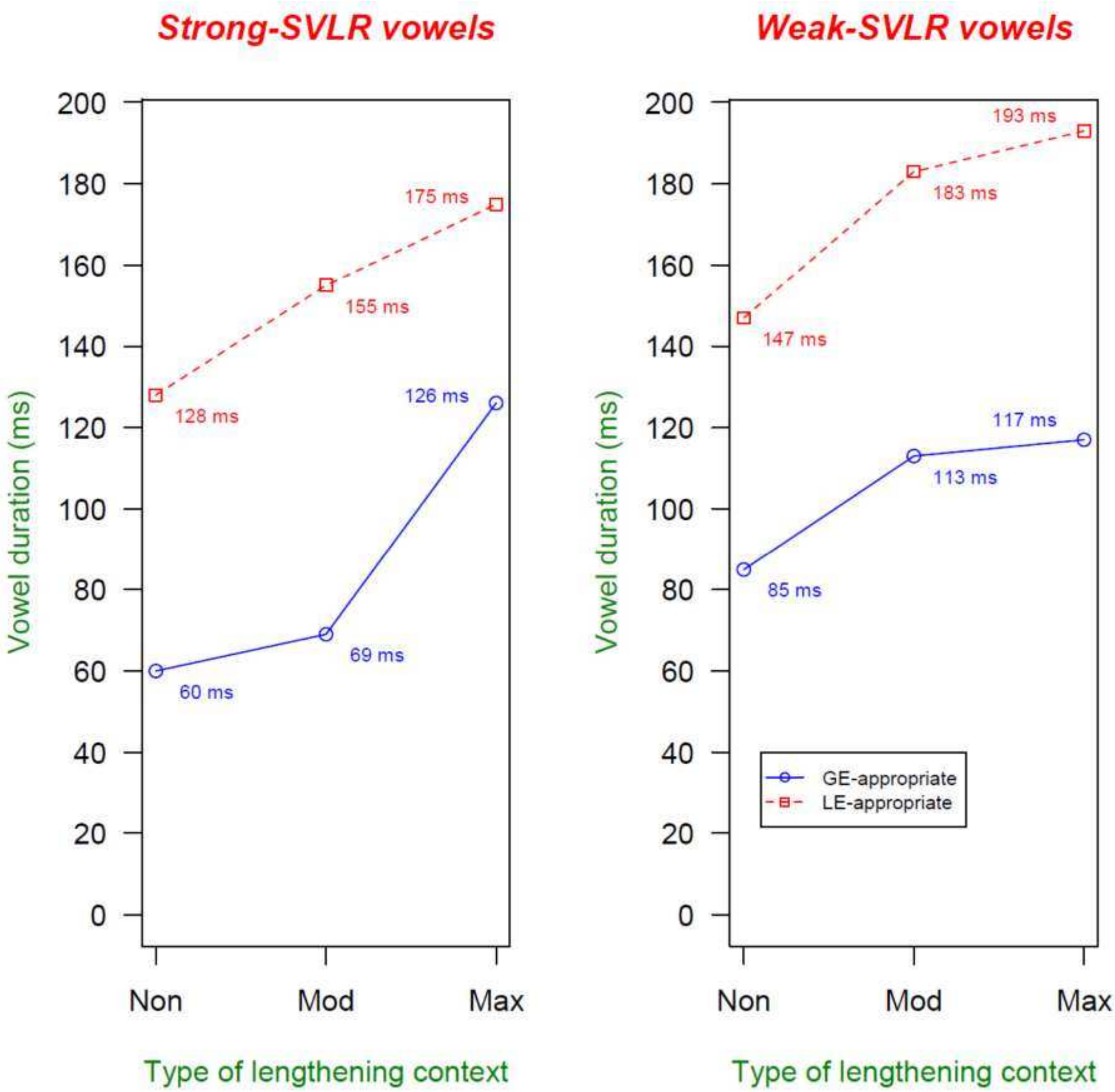

Figure 2. Mean *GE-appropriate* (short) and *LE-appropriate* (long) vowel durations implemented per vowel class and following context. The x-axis shows the type of lengthening context: *non* = *non-lengthening,* (i.e. followed by a voiceless stop or fricative); *mod* = *moderate-lengthening* (i.e. followed by a voiced stop or nasal); *max* = *maximal-lengthening* (i.e. followed by a voiced fricative).

3. Results

3.1. Pre-processing

The dependent variables were spotting rate and reaction time, and each was calculated separately for OVC and OV target words, yielding four dependent variables in total. To calculate spotting rates, the first author coded all spoken responses according to whether the participant had spotted the OVC word, the OV word, a different word or no word. The proportions of OVC and OV responses were calculated separately, each out of the total of 66 experimental trials. For the LE data, a few cases of uncertainty as to the word spoken were presented to a linguistically-trained research assistant in Leeds who was highly familiar with LE, who identified the word spoken. Reaction times were measured from the offset of the OVC embedded word to the time at which the button was pressed. They were log-transformed for statistical analysis, though the text and figures report raw reaction times for ease of interpretation.

Each dependent variable was analysed with linear mixed-effects modelling (Baayen, 2008) using the *lme4* library in R, version 2.15.2. For each variable, the model-fitting procedure was as follows: first a full model was fitted, with the fixed factors Dialect (LE, GE), Vowel duration (LE-appropriate, GE-appropriate), Vowel type (Strong-SVLR, Weak-SVLR), Context (Non-lengthening, Moderate-lengthening, Maximal-lengthening) and all of their interactions, plus control variables Trial, Target log frequency, Gender, Randomisation, and Version. Following Barr, Levy, Scheepers & Tily (2013) the models were fitted with the maximal random effects structure that was justified by the design and allowed the models to converge, crucially including by-subject random slopes for Vowel duration, and by-target random slopes for Dialect.

Predictors that did not significantly contribute to the model were incrementally removed until the simplest model had been found. Model comparison via log-likelihood tests was used to check model fit as predictors were removed. Since this model comparison process does not yield separate significances for main effect terms when the effects are involved in higher-order interactions, main effects and lower-order interactions are only reported when higher-order interactions did not reach significance. Planned comparisons based on z tests were carried out using the *multcomp* package.

3.2. OVC words

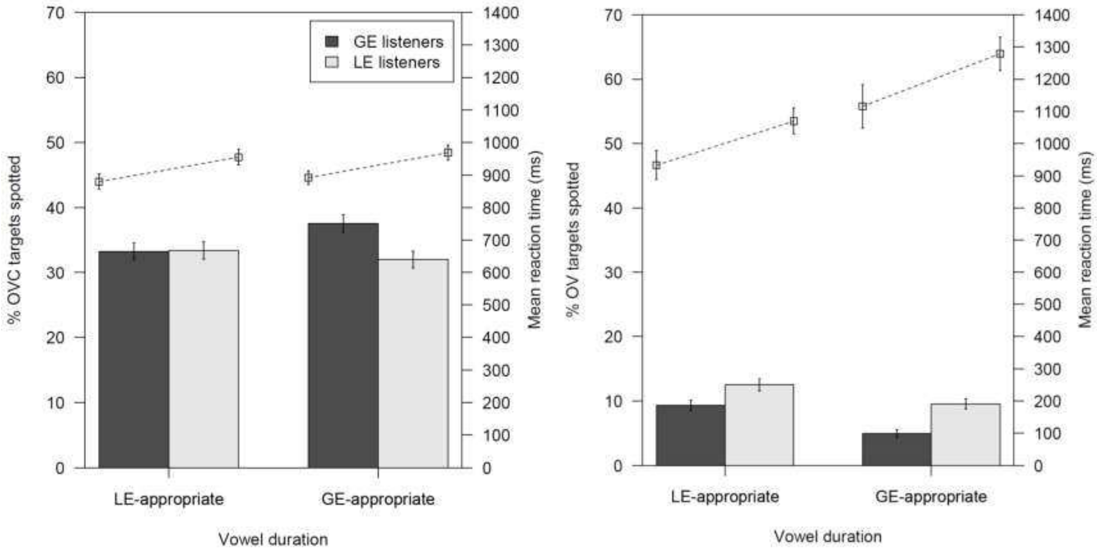

Figure 3. Mean percentage spotting rate (bars) and mean reaction times (lines), by listener accent and vowel duration. Error bars represent 1 standard error. Dark bars and points aligned above them represent GE listeners; light bars and points aligned above them, LE listeners. Left panel shows OVC targets; right panel, OV targets.

Figure 3 (left panel) shows the mean percentage of OVC targets correctly spotted, and the mean reaction times to spot these targets, by accent and vowel duration. Figure 4 breaks these results down by phonological context and vowel type. As predicted, there was a significant interaction of accent with vowel duration (χ^2 (1) = 6.50, p = 0.0108). Planned within-accent comparisons revealed that GE listeners spotted words with GE-appropriate vowels more often than words with LE-appropriate vowels (z = 2.859, p = 0.008) whereas LE listeners were unaffected by vowel duration. Moreover, between-accent comparisons showed that GE listeners spotted words with GE-appropriate vowels more accurately than LE listeners did (z = 2.12, p = 0.034), whereas there was no difference between the listener groups for words with LE-appropriate vowels. There was also a significant interaction of vowel duration, context, and vowel type (χ^2 (2) = 18.29, p < 0.0001). That is, words containing strong-SVLR vowels in non-lengthening and moderate-lengthening contexts, e.g. *freak* in *bezifreakib* and *zoom* in *flizoomip* respectively, were spotted more accurately when their vowel was GE-appropriate (short) than LE-appropriate (long); the difference was 17.2% in non-lengthening contexts (z = 6.758, $p < 0.0001$) and 6.6% in moderate-lengthening contexts (z = 3.076, p = 0.0124). However, vowel duration produced no significant difference in spotting rates for strong-SVLR vowels in maximal-lengthening contexts, or for weak-SVLR vowels in any context. Although the four-way interaction between accent, vowel duration, context, and vowel type did not reach significance, Figure 4 suggests that the advantage for GE-appropriate strong-SVLR vowels emerges in part due to the GE participants' noticeably better performance with these vowels, especially in moderate-lengthening contexts (e.g. *need, zoom*). Increasing frequency of the target word improved the spotting rate (χ^2 (1) = 15.12, $p <$ 0.0001), female participants spotted slightly more target words than males (χ^2 (1) = 5.56, p = 0.0183), and a significant effect of trial shows that spotting rate improved over the course of the experiment (χ^2 (1) = 21.05, $p < 0.0001$).

The analysis of reaction times yielded relatively few significant effects, perhaps because numbers of correct responses were relatively low. There was a significant interaction of accentwith vowel duration and context (χ^2 (2) = 6.00, p = 0.0497). Planned comparisons revealed that LE listeners were slower than GE listeners to spot words containing GE-appropriate vowel variants in non-lengthening contexts, e.g. *freak* in *bezifreakib* (mean difference between LE and GE listeners 219 ms; z = 1.971, p = 0.0487), whereas differential accent-based effects of vowel duration did not occur in other contexts. Comparison of the top left and top right panels of Figure 4 suggests that the slowing of LE listeners' responses to GE-appropriate variants in non-lengthening contexts was limited to the strong-SVLR vowels, /i/ and /u/, whose GE-appropriate variants were (at 60 ms) the shortest in the dataset, and this pattern may underlie the marginally significant interaction of vowel duration, context and vowel type (χ^2 (2) = 5.13, p = 0.0769). Males responded slightly more slowly than females (χ^2 (1) = 3.93, p = 0.0497), and increasing frequency of the target word reduced reaction times (χ^2 (1) = 8.04, p = 0.0046).

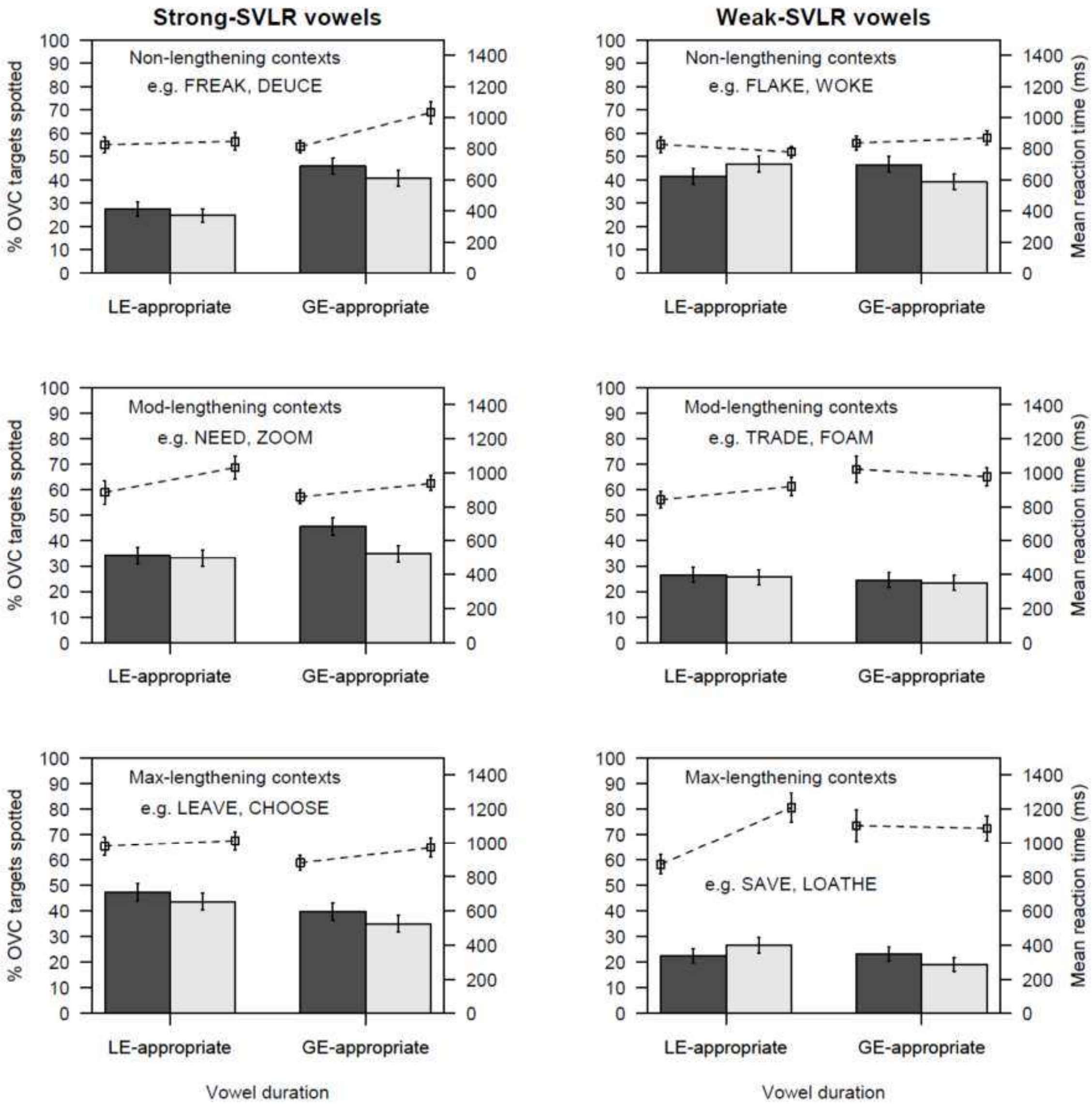

Figure 4. Mean percentage spotting rate (bars) and mean reaction times (lines) to spot OVC target words, by listener accent and vowel duration. Panels are arranged by vowel type (columns) and phonological context (rows). Error bars for RT data represent 1 standard error. Dark bars and points aligned above them represent GE listeners; light bars and points aligned above them, LE listeners.

3.3. OV targets

OV targets (e.g. *fee* in *sifeedizh*) were spotted far less frequently overall than OVC targets (*feed* in *sifeedizh*): 9% vs 34% overall. Figure 3 (right panel) shows the mean percentage of OV targets correctly spotted, and the mean reaction times to spot these targets, by accent and vowel duration. Figure 5 breaks these results down by phonological context and vowel type. There was a significant interaction of vowel duration with context (χ^2 (2) = 30.70, $p <$ 0.0001): hearing an LE-appropriate rather than a GE-appropriate variant increased listeners' likelihood of spotting an OV target in both non-lengthening contexts (15.5% vs 6.5%, z = 5.832, $p < 0.0001$) and moderate-lengthening contexts (13.4% vs 9.4%, z = 2.668, $p <$ 0.0001) but not in maximum-lengthening contexts (3.8% vs 5.8%, ns). The interaction of accent and vowel duration was not significant. However, accent did affect responding in interaction with vowel type (χ^2 (1) = 12.14, $p = 0.0005$): LE listeners spotted substantially significantly more OV targets than GE listeners did for words containing weak-SVLR vowels (10.0% vs 4.5%, z = 4.567, $p < 0.0001$) but only marginally more for words containing strong-SVLR vowels (12.0% vs 9.8%, z = 1.990, $p = 0.0833$). Vowel duration also interacted significantly with vowel type (χ^2 (1) = 4.46, $p = 0.0348$): for strong-SVLR vowels, listeners of both accents spotted OV targets in 14.1% of cases when vowels were LE-appropriate, and only 7.8% of cases when vowels were GE-appropriate (z = 3.224, $p = 0.0025$) but vowel duration did not affect the spotting rate for weak-SVLR vowels (7.8% vs 6.8% for long vs short variants, n.s.). A significant interaction of vowel type with context (χ^2 (2) = 11.13, $p =$ 0.0038) reflects the fact that in non-lengthening contexts, OV targets were more frequently spotted in words with strong-SVLR than weak-SVLR vowels (17.1% vs 4.9%, z = 3.635, $p =$ 0.0008) but this pattern was not found in the other phonological contexts. Increasing frequency of the target word again improved the spotting rate (χ^2 (1) = 15.64, $p < 0.0001$), and the spotting rate improved over the course of the experimental trials (χ^2 (1) = 5.82, $p =$ 0.0158).

Reaction times to OV words were slower overall than to OVC words (923 ms vs 1096 ms). RTs to OV words were significantly faster when listeners heard LE-appropriate (long) vowel variants (1011 ms) than GE-appropriate (short) variants (1223 ms; χ^2 (1) = 10.34, $p =$ 0.0013). RTs were additionally significantly affected by phonological context (χ^2 (2) = 6.03,

$p = 0.049$): responses were faster in non-lengthening contexts than in maximum-lengthening contexts (952 vs 1198 ms; $z = 2.496$, $p = 0.0392$) with moderate-lengthening contexts yielding intermediate RTs (1026 ms). LE listeners tended to respond more slowly than GE listeners (1161 vs 996 ms on average, n.s.).

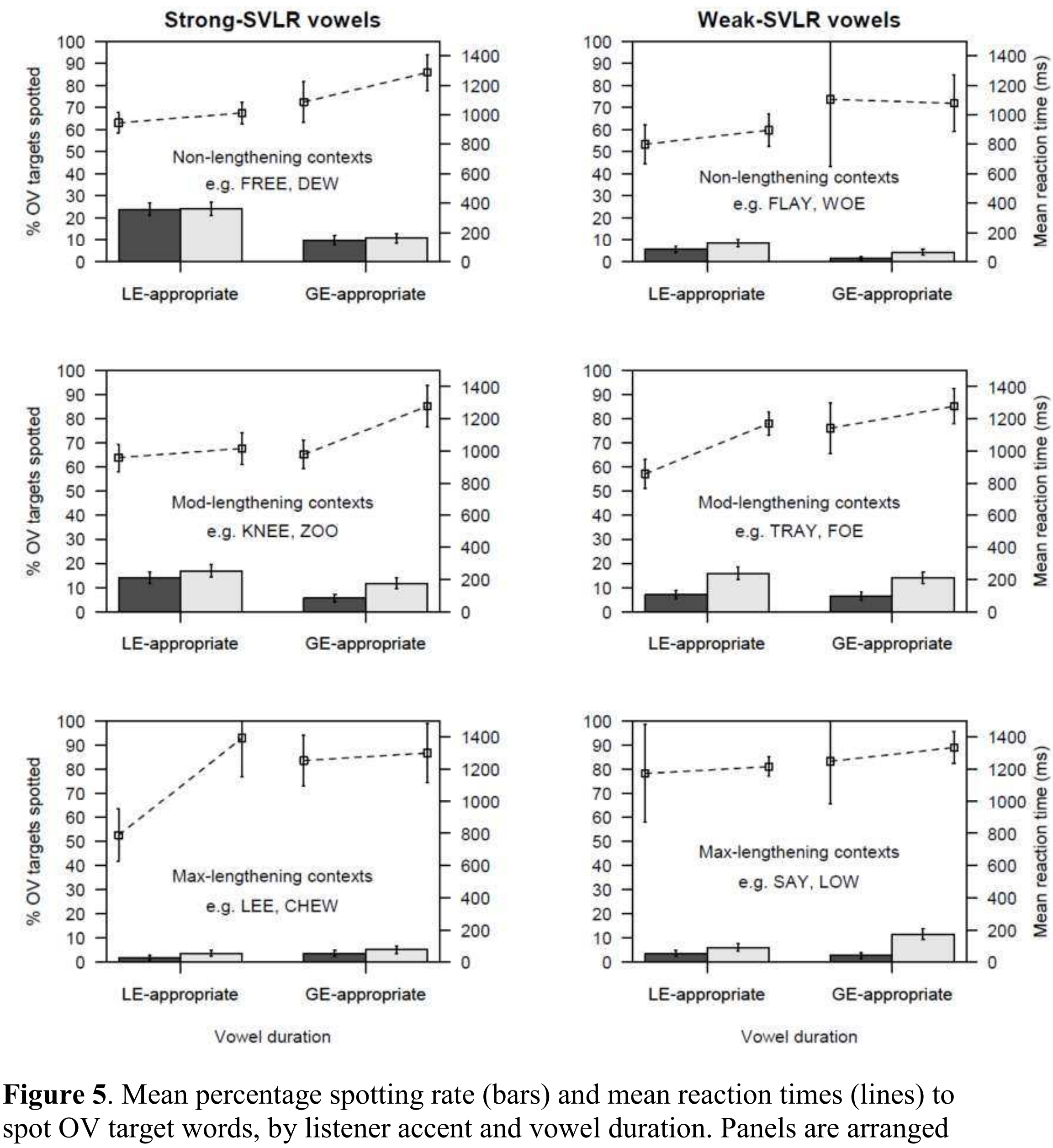

Figure 5. Mean percentage spotting rate (bars) and mean reaction times (lines) to spot OV target words, by listener accent and vowel duration. Panels are arranged by vowel type (columns) and phonological context (rows). Error bars for RT data represent 1 standard error. Dark bars and points aligned above them represent GE listeners; light bars and points aligned above them, LE listeners.

4. Discussion

Previous studies of cross-accent speech comprehension difficulties have either lacked phonetic specificity or have focused on spectral cues to phonemic distinctions. Yet durational properties show extensive cross-accent variation and, crucially, fulfil multiple perceptual functions extending well beyond the level of segmental identity. A better understanding of how accent variation in duration is processed can help to constrain models of spoken word recognition with respect to how such variation affects the time course of lexical competition and to how complex durational dependencies are learned about and represented. The present study used word-spotting to investigate these questions, examining how often and how fast speakers of Glasgow English (GE) and Leeds English (LE) spotted words such as *gloom* or its competitor *glue,* embedded within nonsense sequences such as *pobegloomezh,* according to the duration of the vowel and the phonological context.

Overall, listeners' performance in the word-spotting task was poorer than in many word spotting experiments (around 43% on average when both OV and OVC targets are summed, compared to 55%-80% in other published word-spotting experiments; Norris, McQueen, & Cutler, 1995; Norris, McQueen, Cutler, & Butterfield, 1997; McQueen, 1998). This may reflect a number of factors. First, the frequencies of our target words were rather low, which could not be avoided as the words were selected primarily to satisfy phonetic/phonological constraints. Second, there was always competition between two word candidates present in the same stimulus, which likely makes it harder for either candidate to reach some hypothetical threshold for response. Third, the nonsense words in our experiment were also longer (at 3 or 4 syllables) than in classic studies (typically 2 syllables); this will have made them intrinsically more demanding to process, and also created more opportunities for errors, i.e. a longer nonsense string contains more sites where spurious "words" other than the intended targets could be spotted. (To offset this, the position of the target was to some extent predictable - it was always in the penultimate syllable, and the presence of primary stress was a cue to the target-containing syllable.) The use of an accent other than the participants' own may have lowered spotting rates. The improvement in spotting rates over the course of the experimental trials confirms that the task was difficult, but became easier with experience, so for future work, a longer practice period, perhaps with feedback, might be useful. Nevertheless, despite the relatively low spotting rates, differences between conditions did emerge, which demonstrate subtle yet significant effects of duration on word-spotting, as we now discuss.

The experiment demonstrated, as expected (question 1a), that accent-inappropriate vowel duration affected word-spotting. However, the effects were asymmetrical across the accents. That is, GE listeners were more accurate at spotting words like *gloom* when vowel duration was appropriate for GE (i.e., short) than when it was inappropriate (long), while LE listeners did not respond more accurately to LE-appropriate (long) than to GE-appropriate (short) vowel durations. Nevertheless, LE listeners were less accurate (and in some contexts slower) relative to GE listeners in responding to GE-appropriate variants, whereas GE listeners performed no worse than LE listeners in responding to LE-appropriate variants. We consider first why short, GE-appropriate duration advantaged GE listeners, and second why a discrepancy between the groups of listeners emerged only for short, GE-appropriate vowels.

The difference in GE listeners' rates of spotting OVC words with GE-appropriate versus LE-appropriate vowel durations demonstrates that duration is an important aspect of vowel identity for these listeners. This may be because duration is, in a few word pairs like *brood—brewed*, "quasi-phonemic" in GE (Scobbie & Stuart-Smith, 2008) and such cases appear to be perceived somewhat categorically (Ferragne et al., 2011; Boulenger et al., 2011). LE listeners may pay less attention to the duration of a vowel because they rely more on its quality, and less on duration, to determine segment and word identity (though to test this view would require directly pitting quality and duration cues against each other, which this study did not do).

Nonetheless account must be taken of the fact that GE listeners had an advantage over LE listeners in terms of spotting rate and (in some contexts) also speed, for OVC target words when the vowel was short (i.e. appropriate for the GE accent). This suggests that LE listeners were not exempt from disruption by inappropriate vowel duration, whereas the opposite pattern (an advantage for LE over GE listeners in spotting words with long, LE-appropriate vowels) was not found. Why did a difference between the listener groups emerge when vowel duration was GE-appropriate and not when it was LE-appropriate? A listener-based account and a processing account (not mutually exclusive) can be proposed to explain this finding.

The listener-based explanation relates to familiarity with the patterns of the other accent. We chose the two accents carefully in order to avoid their being differentially familiar, and have no reason to believe that GE listeners were in fact specifically familiar with LE speech. What is difficult to eliminate, however, is an accent's similarity to other accents with which participants might be familiar. Specifically, LE has some similarity in its durational patterning with Standard Southern British English (SSBE) and General American (GenAm): in all these varieties vowels are substantially longer before voiced stops, nasals, and voiced fricatives than before voiceless stops. Familiarity with standard accents, e.g. via media exposure, has been often argued to underlie asymmetries in cross-dialect speech perception (Floccia et al, 2006, Adank et al, 2009; Smith & Knight, under revision). Possibly, familiarity with SSBE or even GenAm durational patterning might have helped GE listeners to process LE-appropriate long vowels, whereas LE listeners are unlikely to have been exposed to SVLR-like durational patterning. The idea that the cross-accent prevalence of a phonetic pattern affects its processing is supported by Brunellière, Dufour, Nguyen and Frauenfelder (2009) who showed, for French, a difference in the stability of two phonemic contrasts (as indexed electrophysiologically by MMN and P200 components) for Swiss listeners. A contrast that occurs patchily across accents of French was less stable in electrophysiological terms than a contrast that is common across all French accents, even though the Swiss listeners use both contrasts in their own accent.

The processing account of the asymmetry relates to the differential effects of unexpectedly short vs unexpectedly long vowels on the processing of temporally-organized information. Unexpectedly short vowel variants (e.g., for LE listeners, a GE-appropriate /i/ in *freak* lasting 60 ms) may surprise the perceptual system by ending at a point when activation of words containing the vowels has had little time to build up, whereas when a variant is unexpectedly long (e.g., for GE listeners, an LE-appropriate /i/ in *freak* lasting 128 ms) the perceptual system has time to integrate the unexpected information (cf. Hawkins & Nguyen, 2004). We may compare the effect of short duration in this regard to that of fast rate: Schwab, Miller, Grosjean & Mondini (2008) found word segmentation to be poorer for talkers with fast than slow rates. Differential effects of too-short versus too-long vowel duration—and thus too-early versus too-late arrival of information about the following consonant—also find a natural explanation in the context of dynamic attending theory (Large & Jones, 1999) according to which attention is (quasi-)cyclically distributed, with the consequence that information arriving early in an attentional cycle receives less attention, whereas information arriving late in an attentional cycle receives additional attention (cf. McAuley & Fromboluti, 2014). Processing difficulties associated with short duration may thus explain why LE listeners performed worse than GE listeners with short, GE-appropriate vowels; we will return later in the discussion to attempt to reconcile the apparent contradiction of this aspect of our findings, with the fact that LE listeners performed no better on long, LE-appropriate than short, GE-appropriate vowels.

Turning to the dependence of the results on the details of accent-specific contextual conditioning of duration (question 1b), we found a mixed picture. The clearest divergence between the accents had been predicted to occur for strong-SVLR vowels before voiced stops and nasals (e.g. *gloom* in *pobegloomezh*), i.e. in the context that conditions a much shorter duration in GE than LE (Figure 1). While such a divergence was present numerically for spotting rate, it did not emerge statistically in the form of an expected significant interaction of accent, context, and vowel duration. Instead, LE listeners were less accurate across the board than GE listeners in spotting words with GE-appropriate vowel duration. Conversely we saw more divergence than expected in reaction times in the non-lengthening context, which is the context where vowels are at their shortest for both GE and LE, but in absolute terms are even shorter in GE than LE. The results strengthen the argument that the aspect of GE vowels that is perceptually most disruptive to listeners who speak with other accents is their overall shortness. Listeners appear to be simply less sensitive to the finer details of the contextual conditioning of the SVLR. This would be in line with data suggesting that SVLR-related patterns can be difficult to acquire and susceptible to sound change (Gregg 1973; Kerswill 1996; Labov 1994; Trudgill 1986; Hewlett, Mathews & Scobbie 1999), as well as with limits on the learnability of phonological patterns in artificial languages (Skoruppa and Peperkamp 2011). Nonetheless, further research is required to establish whether there are circumstances under which listeners would be more sensitive to subtle aspects of the contextual conditioning of vowel duration.

The results for spotting OV words did not fit the pattern we predicted (question 2), namely that Glasgow listeners would interpret long, LE-appropriate vowel durations as cueing wordfinality and would spot more OV target words than LE listeners. Instead LE listeners were slightly but significantly more likely than GE listeners to insert word boundaries after vowels. This surprising finding should be considered in the context of several other striking aspects of the OV results. First, the rates of spotting OV words were very low overall compared to those for OVC words: that is, listeners showed a bias to respond (if they spotted a word at all) with the longest available word. This bias is consistent with the conclusion of Pitt & Samuel (2006) that longer words generate greater activation which is available both earlier and for a longer time than is the case for shorter words (cf. findings of Davis, Gaskell, & Marslen-Wilson, 2002, and predictions of neural network models by Grossberg & Stone, 1986; Frauenfelder & Peeters, 1990). Second, there was an interaction between vowel duration and context: GE-appropriate vowels caused significantly fewer OV words to be spotted than LE-appropriate vowels did in non-lengthening and moderate-lengthening contexts, but not in maximal-lengthening contexts. Though not specifically predicted, this interaction makes sense on both phonetic grounds and processing grounds. Phonetically, context-specific expectations (regardless of accent) will favour longer vowels in the maximal-lengthening context, and shorter ones in the moderate-lengthening and non-lengthening contexts: thus long vowels in the latter contexts will be more readily interpreted as word-final. In processing terms, GE-appropriate short duration likely enhances the intrinsic bias to respond with a longer OVC word (e.g. *freak*), because it allows less time for the activation of the OV competitor (e.g. *free*) to build up before the following consonant arrives and boosts the activation of the OVC word (cf. Pitt & Samuel, 2006).

Integrating the above ideas, we can return to the asymmetry in our results whereby GE listeners spot GE-appropriate (short) variants of OVC words more accurately than LE-appropriate (long) variants, while LE listeners do not show the corresponding advantage for LE-appropriate over GE-appropriate variants, but *do* process GE-appropriate variants poorly compared to GE listeners, as well as showing a greater propensity to spot OV embedded words. Our best account of these findings, taken together, runs as follows. GE listeners process GE-appropriate variants of OVC words (e.g. *freak*) relatively accurately for two reasons: 1) their short duration is appropriate for their accent, and 2) their short duration enhances the intrinsic bias to respond with an OVC word. They process LE-appropriate variants of OVC words relatively more poorly than short-vowel variants both because long vowels constitute less good a match for their accent-specific expectations, and because their greater duration also leads to increased competition between OVC and OV targets. LE listeners show a disadvantage relative to GE listeners for GE-appropriate variants because these are not appropriate for their accent and end disruptively early relative to expectations. Nevertheless, LE listeners fail to show the predicted advantage in spotting OVC words with LE-appropriate over GE-appropriate vowels because when vowel duration is long, OVC words undergo increased competition with OV words since the activation of the latter has more time to develop before the final consonant begins. Supporting this view, when combined rates of word-spotting are calculated (i.e., the sum of OVC and OV target spotting rates) LE listeners show greater likelihood of spotting *one* of the targets when vowel duration is LE-appropriate (45.7%) than GE-appropriate (41.5%), whereas GE listeners show equivalent spotting rates for LE-appropriate and GE-appropriate durations (42.6% and 42.5% respectively).

In this experiment, we took a novel methodological approach to isolating vowel duration as an accent-specific perceptual cue by using base stimuli spoken by a phonetician to have segmental qualities intermediate between the two accents. The method was a qualified success, and can be recommended to other researchers. It enabled the creation of a large number of distinct stimuli without laborious hand-crafting of synthetic stimuli; the resulting stimuli sounded natural, and were perceived as no more Leeds-like by LE listeners than they were Glasgow-like by GE listeners. Not all intermediate tokens were equally effective, however; our impression was that the weak-SVLR vowels were further off-target than the strong-SVLR vowels, with a few tokens of /o/, in particular, giving rise to vowel misidentifications when heard in their GE-appropriate variants (e.g., *bloke* heard as *block, goad* as *God* or *woke* as *walk*). This may have contributed to our finding stronger effects of context for strong-SVLR than weak-SVLR vowels. Imprecision of match between the qualities of the stimuli and listeners' accent-specific expectations may also have restricted the extent to which we could detect subtler influences of SVLR contextual conditioning. Miller et al. (2011) investigated accent-specific vowel duration perception by applying a durational manipulation to two sets of base stimuli, spoken in Swiss and Parisian accents of French, and found slight differences according to which set of base stimuli was used. In particular, Parisian French listeners, whose accent features a very small durational difference between

/o/ and /ɔ/, were only sensitive to the durational manipulation when base stimuli were spoken in their own accent. While our method was valuable in isolating the role of timing cues alone, for the future it will also be useful to manipulate vowel timing in stimuli spoken with natural base accents. By encouraging listeners to listen in different accent “modes”, we may be able to assess subtle effects involving differences in the weighting of timing and spectral cues across accents.

Our findings, demonstrating effects of accent background on the role of vowel duration in word recognition, are difficult to accommodate within purely abstractionist models of recognition (Halle, 1985). However, the contextual sensitivities that we have demonstrated also challenge exemplar models that are based solely around exemplar representation of segmental categories and/or words (Goldinger, 1996, 1998; Johnson, 1997, 2006; Lachs, McMichael, & Pisoni, 2003; Pisoni, 1997). They are more readily compatible, broadly speaking, with any model of speech processing according to which listeners’ knowledge about phonetic categories is assumed to reflect both phonological structure and specific experience, e.g. hybrid abstract-exemplar accounts (e.g., Pierrehumbert, 2006; Walsh, Möbius, Wade, and Schütze, 2010; Hawkins & Smith, 2001; Hawkins, 2003, 2010), though such models are not currently specified in sufficient detail as to make competing and testable claims about accent variation in vowel duration. The present study contributes to the evolving picture of the role of experience in guiding lexical segmentation and recognition, by focusing on the effects of specific, phonetically detailed cues that differ across accents.

Acknowledgements

This research was funded by the Economic and Social Research Council, grant reference RES-061-25-0302 to Dr R. Smith. The funder had no involvement in study design, data collection, analysis or interpretation of data, writing, or the decision to submit the article for publication. We gratefully acknowledge Mark J. Jones for producing the base stimuli, and Natasha Rust and Ewa Wanat for assistance with data collection.

References

Abercrombie, D. (1979). The accents of standard English in Scotland. In A. J. Aitken &. T. McArthur. (Eds.), *Languages of Scotland* (pp. 68-84). Edinburgh: Chambers.

Adank, P., Smits, R., & van Hout, R. (2004). A comparison of vowel normalization procedures for language variation research. *Journal of the Acoustical Society of America, 116,* 3099-3107.

Adank, P., Evans, B. G., Stuart-Smith, J., & Scott, S. K. (2009). Comprehension of familiar and unfamiliar native accents under adverse listening conditions. *Journal of Experimental Psychology: Human Perception and Performance, 35*(2), 520-529.

Agutter, A. (1988). The not-so-Scottish Vowel Length Rule. In Anderson, J. M. & MacLeod, N. (eds.), *Edinburgh studies in the English language* (pp. 120-132). Edinburgh: John Donald.

Ainsworth, W.A. (1972). Duration as a cue in the recognition of synthetic vowels. *Journal of the Acoustical Society of America, 51,* 648-651.

Aitken, A. J. (1981). The Scottish Vowel Length Rule. In Benskin, M., & Samuels, M. L. (eds.), *So meny people longages and tonges: Philological essays in Scots and Mediaeval English presented to Angus McIntosh* (pp. 131-157). Edinburgh: Middle English Dialect Project.

Baayen, R. H. (2008). *Analysing linguistic data: A practical introduction to statistics.* Cambridge: Cambridge University Press.

Baayen, R., H., Piepenbrock, R. & Gulikers, L. (1985) The CELEX Lexical Database (CD-ROM). Linguistic Data Consortium, University of Pennsylvania, Philadelphia, PA.

Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language, 68*, 255-278.

Beckman, M. E., Edwards, J. (1990). Lengthening and shortening and the nature of prosodic constituency. In J. Kingston and M. E. Beckman (eds.), *Laboratory Phonology I.* (pp. 152-178). Cambridge: Cambridge University Press.

Berkovits, R. (1994). Durational effects in final lengthening, gapping and contrastive stress. *Language and Speech, 37,* 237-250.

Boersma, Paul & Weenink, David (2010). Praat: doing phonetics by computer [Computer program]. Version 5.1.37, retrieved 24 June 2010 from http://www.praat.org/.

Boulenger, V., Ferragne, E., Bedoin, N. & Pellegrino, F. (2011). Derived contrasts in Scottish English: an EEG study. *Proceedings of the 17th International Congress of Phonetic Sciences,* Hong Kong, August 17-21 2011 (pp. 352-355).

Brunellière, A., Dufour, S., Nguyen, N., & Frauenfelder, U.H. (2009). Behavioral and electrophysiological evidence for the impact of regional variation on phoneme perception. *Cognition, 111,* 390-396.

Brunellière, A., Dufour, S., & Nguyen, N. (2011). Regional differences in the listener's phonemic inventory affect semantic processing: A mismatch negativity (MMN) study. *Brain and Language, 117,* 45-51.

Cho, T., McQueen, J. M., & Cox, E. A. (2007). Prosodically driven phonetic detail in speech processing: The case of domain-initial strengthening in English. *Journal of Phonetics*, *35,* 210-243.

Christophe, A., Peperkamp, S., Pallier, C., Block, E., & Mehler, J. (2004). Phonological phrase boundaries constrain lexical access: I. Adult data. *Journal of Memory and Language,* 51*,* 523-547.

Clopper, C.G., & Bradlow, A. (2008). Perception of dialect variation in noise: Intelligibility and classification. *Language and Speech, 51,* 175-198.

Clopper, C. G., Pierrehumbert, J. B., & Tamati, T. N. (2010). Lexical neighborhoods and phonological confusability in cross-dialect word recognition in noise. *Laboratory Phonology, 1*, 65-92.

Cutler, A. and Norris, D. (1988). The role of strong syllables in segmentation for lexical access. *Journal of Experimental Psychology: Human Perception and Performance, 14*, 113-121.

Dahan, D., & Mead, R. L. (2010). Context-conditioned generalization in adaptation to distorted speech. *Journal of Experimental Psychology: Human Perception and Performance, 36,* 704-728.

Davis, M. H., Marslen-Wilson, W., & Gaskell, M. G. (2002). Leading up the lexical garden-path: segmentation and ambiguity in spoken word recognition. *Journal of Experimental Psychology: Human Perception and Performance, 28,* 218-244.

Dilley, L. C. & McAuley, J. D. (2008). Distal prosodic context affects word segmentation and lexical processing. *Journal of Memory and Language,* 59, 294-311.

Dufour, S., Nguyen, N., and Frauenfelder, U. H. (2007). The perception of phonemic contrasts in a non-native dialect. *Journal of the Acoustical Society of America, 121*, EL131–EL136.

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.

Ferragne, E., Bedoin, N., Boulenger, V. & Pellegrino, F. (2011). The perception of a derived contrast in Scottish English. *Proceedings of the 17th International Congress of Phonetic Sciences,* Hong Kong, August 17-21 2011 (pp. 667-670).

Finley, S., & Badecker, W. (2009). Artificial language learning and feature-based generalization. *Journal of Memory and Language*, *61*, 423–437.

Fischer, R. M. & Ohde, R. (1990). Spectral and duration properties of front vowels as cues to final stop-consonant voicing. *Journal of the Acoustical Society of America, 88,* 1250-1259.

Fletcher, J. (2010). The prosody of speech: Timing and rhythm. In Hardcastle, W. J., Laver, J., and Gibbon, F. (eds.) *The handbook of phonetic sciences* (2nd edition). Oxford: Blackwell. 523-602.

Floccia, C., Butler, J., Goslin, J., & Ellis, L. (2009). Regional and foreign accent processing in English: Can listeners adapt? *Journal of Psycholinguistic Research, 38(4*), 379-412.

Floccia, C., Goslin, J., Girard, F., & Konopczynski, G. (2006). Does a regional accent perturb speech processing? *Journal of Experimental Psychology: Human Perception and Performance, 32* (5), 1276-1293.

Forster, K.I. & Forster, J.C. (2003). DMDX: A Windows display program with millisecond accuracy. *Behavior Research Methods, Instruments, & Computers, 35* (1), 116-124.

Frauenfelder, U. H., & Peeters, G. (1990). Lexical segmentation in TRACE: An exercise in simulation. In G. T. M. Altmann (Ed.), *Cognitive models of speech processing: Psycholinguistic and computational perspectives* (pp. 50–86). Cambridge, MA: MIT Press.

Goldinger, S. D. (1996). Words and voices: Episodic traces in spoken word identification and recognition memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 22,* 1166-1183.

Goldinger, S. D. (1998). Echoes of echoes? An episodic theory of lexical access. *Psychological Review, 105,* 251-279.

Grossberg, S. & Stone, G. (1986). Neural dynamics of word recognition and recall: Attentional priming, learning, and resonance. *Psychological Review, 93*, 46-74.

Hall, K. C. (2003). The perception of contrast and the Scottish Vowel Length Rule. Unpublished B.A. dissertation, University of North Carolina, Chapel Hill.

Halle, M. (1985). Speculation about the representation of words in memory. In Fromkin, V. (ed.), *Phonetic Linguistics.* New York: Academic Press. 101-114.

Hawkins, S. (2003). Roles and representations of systematic fine phonetic detail in speech understanding. *Journal of Phonetics, 31,* 373-405.

Hawkins, S. (2010). Phonetic variation as communicative system: Perception of the particular and the abstract. In C. Fougeron, B. Kühnert, M. d'Imperio, & N. Vallée (eds.), *Laboratory Phonology 10: Variability, Phonetic Detail and Phonological Representation*. Berlin: Mouton de Gruyter. 479-510.

Hawkins, S., & Nguyen, N. (2004). Effects on word recognition of syllable-onset cues to syllable-coda voicing. In Local, J., Ogden, R., & Temple, R. (eds.), *Phonetic Interpretation: Papers in Laboratory Phonology VI.* 38-57.

Hawkins, S., & Smith, R. H. (2001). Polysp: a polysystemic, phonetically-rich approach to speech understanding. *Italian Journal of Linguistics-Rivista di Linguistica, 13,* 99-188.

Hewlett, N., Mathews, B., & Scobbie, J. M. (1999). Vowel duration in Scottish English speaking children. *Proceedings of the XIVth International Congress of Phonetic Sciences.* University of California: Berkeley.

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

Jacewicz, E., & Fox, R. A. (2012). The effects of cross-generational and cross-dialectal variation on vowel identification and classification. *Journal of the Acoustical Society of America, 131,* 1413-1433.

Johnson, K. (1997). Speech perception without speaker normalization: An exemplar model. In K. Johnson, & J. W. Mullennix (Eds.), *Talker variability in speech processing* (pp.145-165). San Diego: Academic Press.

Johnson, K. (2006). Resonance in an exemplar-based lexicon: The emergence of social identity and phonology. *Journal of Phonetics, 34,* 485-499.

Johnston, P. A. (1997). Older Scots phonology and its regional variation. In Jones, C. (ed.), *The Edinburgh history of the Scots language* (pp. 47-111). Edinburgh: Edinburgh University Press.

Keating, P. A. (1985). Universal phonetics and the organization of grammars. In Fromkin, V. (ed.), *Phonetic linguistics: Essays in honor of Peter Ladefoged* (pp. 115-132). Orlando: Academic Press.

Kemps, R. J. J. K., Ernestus, M., Schreuder, R., & Baayen, R. H. (2005). Prosodic cues for morphological complexity: The case of Dutch plural nouns. *Memory & Cognition,* 33, 430–446.

Kim, S. and Cho, T. (2009). The use of phrase-level prosodic information in lexical segmentation: Evidence from word-spotting experiments in Korean. *Journal of the Acoustical Society of America*, 125, 3373-3386.

Labov, W. and Ash, S. (1997). Understanding Birmingham. In Bernstein, C., Nunnally, T., & Sabino, R. (eds.), *Language variety in the South revisited* (pp. 508-573). Tuscaloosa: University of Alabama Press.

Lachs, L., McMichael, K., & Pisoni, D. B. (2003). Speech perception and implicit memory: Evidence for detailed episodic encoding. In J. S. Bowers, & C. J. Marsolek (Eds.), *Rethinking implicit memory* (pp. 215-235). Oxford: Oxford University Press.

Large, E. W., & Riess Jones, M. (1999). The dynamics of attending: How people track time-varying events. *Psychological Review, 106*, 119-159.

Lehiste, I. (1977). Isochrony reconsidered. *Journal of Phonetics, 5,* 253-263.

Lehiste, I., Olive, J. P. & Streeter, L. (1972). Role of duration in disambiguating syntactically ambiguous sentences. *Journal of the Acoustical Society of America, 60*, 1199–1202.

Lobanov, B. M. (1971). Classification of Russian vowels spoken by different speakers. *Journal of the Acoustical Society of America, 49*, 606–608.

McAuley, J. D. & Kim, E. (in preparation). Attentional entrainment and perceived event duration. Manuscript in preparation.

McKenna, G. E. (1988). Vowel duration in the Standard English of Scotland. Unpublished M. Litt. dissertation, University of Edinburgh.

McQueen, J.M. (1998). Segmentation of continuous speech using phonotactics. *Journal of Memory and Language, 39*, 21-46.

Maddieson, I. (1985). Phonetic cues to syllabification. In Fromkin, V. A. (ed.), *Phonetic linguistics: Essays in honor of Peter Ladefoged* (pp. 203-221). Orlando: Academic.

Magnuson, J. S., Dixon, J. A., Tanenhaus, M. K., & Aslin, R. N. (2007). The dynamics of lexical competition during spoken word recognition. *Cognitive Science, 31,* 233-256.

Magnuson, J. S. & Nusbaum, H. C. (2007). Acoustic differences, listener expectations, and the perceptual accommodation of talker variability. *Journal of Experimental Psychology: Human Perception and Performance, 33*, 391-409.

Mermelstein, P. (1978). On the relationship between vowel and consonant identification when cued by the same acoustic information. *Perception & Psychophysics, 23,* 331-336.

Miller, J. L. & Grosjean, F. (1997). Dialect effects in vowel perception: The role of temporal information in French. *Language and Speech, 40,* 277-288.

Miller, J.L., Mondini, M., Grosjean, F. and Dommergues, J.-Y. (2011). Dialect effects in speech perception: The role of vowel duration in Parisian French and Swiss French. *Language and Speech, 54,* 467-485.

Milroy, J. (1995). Investigating the Scottish Vowel Length Rule in a Northumbrian dialect. *Newcastle and Durham Working Papers in Linguistics, 3,* 187-196.

Nathan, L., & Wells, B. (2001). Can children with speech difficulties process an unfamiliar accent? *Applied Psycholinguistics, 22*(3), 343-361.

Nathan, L., Wells, B., & Donlan, C. (1998). Children's comprehension of unfamiliar regional accents: A preliminary investigation. *Journal of Child Language, 25*(2), 343-365.

Nooteboom, S. G., & Doodeman, G. J. N. (1980). Production and perception of vowel length in spoken sentences. *Journal of the Acoustical Society of America, 67,* 276-287.

Norris, D., McQueen, J. M., & Cutler, A. (1995). Competition and segmentation in spoken-word recognition. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 21,* 1209-1228.

Norris, D., McQueen, J. M., Cutler, A., & Butterfield, S. (1997). The possible-word constraint in the segmentation of continuous speech. *Cognitive Psychology, 34,* 191-243.

Nygaard, L. C., & Pisoni, D. B. (1998). Talker-specific learning in speech perception. *Perception and Psychophysics, 60,* 355-376.

Peperkamp, S., & Dupoux, E. (2007). Learning the mapping from surface to underlying representations in an artificial language. In J. Cole, & J. Hualde (eds.), *Laboratory Phonology 9* (pp. 315–338). Berlin: Mouton de Gruyter.

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

Pierrehumbert, Janet (2003). Probabilistic phonology: Discrimination and robustness. In Rens Bod, Jennifer Hay, and Stephanie Jannedy, editors, Probabilistic Linguistics, pages 97-138. MIT Press, Malden MA, 2003.

Pierrehumbert, J. (2006). The next toolkit. *Journal of Phonetics, 34,* 516-530.

Pisoni, D. B. (1997). Some thoughts on "normalization" in speech perception. In K. Johnson, & J. W. Mullennix (Eds.), *Talker variability in speech processing* (pp. 9–32). San Diego: Academic Press.

Pitt, M.A., & Samuel, A.G. (2006) Word length and lexical activation: Longer is better. *Journal of Experimental Psychology: Human Perception and Performance, 32,* 1120-1135.

Salverda, A. P., Dahan, D., & McQueen, J. M. (2003). The role of prosodic boundaries in the resolution of lexical embedding in speech comprehension. *Cognition, 90,* 51-89.

Scobbie, J. M., Hewlett, N., & Turk, A. (1999). Standard English in Edinburgh and Glasgow: the Scottish Vowel Length Rule revisited. In Foulkes, P. & Docherty, G. (eds.), *Urban voices: Accent studies in the British Isles* (pp. 230-245). London: Hodder.

Scobbie, J.M., & Stuart-Smith, J. (2008). Quasi-phonemic contrast and the fuzzy inventory: examples from Scottish English. In Avery, P., Dresher, B.E., & Rice, K. (eds.), *Contrast in phonology: Theory, perception, acquisition* (pp. 87-113). Berlin: Mouton de Gruyter.

Scott, D. R. (1982). Duration as a cue to the perceptibility of a phrase boundary. *Journal of the Acoustical Society of America, 71,* 996-1007.

Scott, D. R. & Cutler, A. (1984). Segmental phonology and the perception of syntactic structure. *Journal of Verbal Learning and Verbal Behavior, 23,* 450-466.

Schwab, S., Miller, J. L., Grosjean, F. & Mondini, M. (2008). Effect of speaking rate on the identification of word boundaries. *Phonetica, 65,* 173-186.

Skoruppa, K., and Peperkamp, S. (2011). Adaptation to novel accents: Feature-based learning of context-sensitive phonological regularities. *Cognitive Science, 35,* 348-366.

Smith, R. and Hawkins, S. (2000). Allophonic influences on word-spotting experiments. In Cutler, A., McQueen, J.M. and Zondervan, R. (eds.), *Proceedings of the Workshop on Spoken Word Access Processes (SWAP),* Nijmegen, The Netherlands. 139-142.

Smith, R., & Hawkins, S. (2012). Production and perception of speaker-specific phonetic detail at word boundaries. *Journal of Phonetics, 40*, 213-233.

Smith, R., Knight, R.-A., Holmes-Elliott, S., & Pettinato, M. (under revision). Cross-accent intelligibility of speech in noise: Long-term familiarity and short-term familiarization. Under revision for *Quarterly Journal of Experimental Psychology.*

Sóskuthy, M. (2015). Formant Editor: Software for editing dynamic formant measurements. Version 0.8.2, retrieved 08 September 2015 from https://github.com/soskuthy/formant_edit

Strange, W. (1989). Dynamic specification of coarticulated vowels spoken in sentence context. *Journal of the Acoustical Society of America, 85,* 2135-2153.

Stuart-Smith, J. (1999). Glasgow: Accent and voice quality. In P. Foulkes, & G. Docherty (eds.), *Urban voices: Accent studies in the British Isles* (pp. 203-222). London: Arnold.

Stuart-Smith, J. (2003). The phonology of modern urban Scots. In J. Corbett, J. D. McClure & J. Stuart-Smith (eds.), *The Edinburgh Companion to Scots* (pp. 110-137). Edinburgh: Edinburgh University Press.

Stuart-Smith, J. (2004). Scottish English: phonology. In B. Kortmann, K. Burridge, E.W. Schneider, R. Mesthrie, & C. Upton (eds.) *A Handbook of Varieties of English: 1: Phonology* (pp. 47-67). Berlin: Mouton de Gruyter.

Sumner, M., & Samuel, A. G. (2009). The effect of experience on the perception and representation of dialect variants. *Journal of Memory and Language, 60* (4), 487-501.

Thiessen, E. D., & Saffran, J. R. (2004). Spectral tilt as a cue to word segmentation in infancy and adulthood. *Perception and Psychophysics, 66,* 779-791.

Walsh, M., Möbius, B., Wade, T., & Schütze, H. (2010). Multilevel exemplar theory. *Cognitive Science, 34,* 537-582.

Watt, D. & Ingham, C. (2000). Durational evidence of the Scottish Vowel Length Rule in Berwick English. In Nelson, D. & P. Foulkes (eds) Leeds Working Papers in Linguistics 8, pp. 205-228.

Welby, P. (2009). The role of early fundamental frequency rises and elbows in French word segmentation. *Speech Communication, 49*, 28-48.

Wells, J. C. (1982). *Accents of English*. Cambridge: Cambridge University Press.

Whalen, D. H. (1989). Vowel and consonant judgements are not independent when cued by the same information. *Perception & Psychophysics, 46,* 284-292.

Wightman, C.W., Shattuck-Hufnagel, S., Ostendorf, M., Price, P.J. (1992). Segmental Durations in the Vicinity of Prosodic Phrase Boundaries. *Journal of the Acoustical Society of America 91(3),* 1707-1717.

Appendix Nonsense sequences and their embedded experimental target words. Targets' log frequencies (per million) are from the combined spoken and written COBUILD databases in CELEX (Baayen et al., 1995). For homophones, the sum of the homophones' frequencies is given.

	Strong-SVLR vowels			**Weak-SVLR vowels**		
	Nonsense sequence	**OVC target**	**OV target**	**Nonsense sequence**	**OVC target**	**OV target**
Non-lengthening contexts (voiceless stops or fricatives)	ʒanəˈdjusəθ	deuce (0)	dew/due (2.0)	lɒʃəˈblokəf	bloke (0.8)	blow (1.4)
	pɪˈdrupətʃ	droop (0)	drew (1.2)	ləˈbotɪs	boat (1.8)	bow (1.0)
	ʒɛkwəˈflukəb	fluke (0)	flew (0.8)	taʒɪˈgrosɪm	gross (1.3)	grow (1.4)
	tʃəˈnjutɪs	newt (0)	(k)new (2.1)	jɪˈmopɪb	mope (0)	mow (0)
	ləˈpjukɪv	puke (0)	pew (0.3)	fɛbləˈsopəʒ	soap (1.3)	sew/so (3.5)
	dɪˈtrusədʒ	truce (0.5)	true (2.4)	zʌgɪˈθrotəg	throat (1.6)	throw (1.2)
	gavɪˈflisəʒ	fleece (0)	flee/flea (0.5)	pəˈwokəʃ	woke (0.8)	woe (0.5)
	bɛzɪˈfrikɪb	freak (0.7)	free (2.3)	kɛsəˈflekɪdʒ	flake (0.3)	flay (0)
	ʃɪkəˈpisəg	peace (2.0)	pea/pee (0.5)	ʒɪˈgrepətʃ	grape (0.3)	grey (1.9)
	swəˈplitɪs	pleat (0)	plea (0.9)	dʒapləˈhetɪv	hate (1.4)	hay/hey (1.5)
	ʒɪˈtritɪb	treat (1.1)	tree (1.9)	nɪˈstretɪʒ	straight (2.1)	stray (0.7)
		mean 0.39	***mean 1.34***		***mean 1.06***	***mean 1.19***
		(sd 0.64)	***(sd 0.80)***		***(sd 0.67)***	***(sd 0.98)***
Moderate-lengthening contexts (voiced stops or nasals)	zɪˈkjubəf	cube (0.7)	cue/queue (1.2)	kɪˈgodɪʒ	goad (0)	go (2.5)
	dʒʌməˈbidəf	bead (0.3)	bee/be (3.8)	vafɪˈrobɪð	robe (1.0)	row/roe (1.4)
	sɪˈfidɪʒ	feed (1.4)	fee (1.1)	fɪˈdʒedək	jade (0.3)	J/jay (1.8)
	ʃɛtwəˈnidəb	need (2.4)	knee (1.5)	sɒʃəˈplegɪb	plague (0.8)	play (2.2)
	vɛʃəˈplidɪdʒ	plead (0)	plea (0.9)	lʌməˈtredɪʒ	trade (2.2)	tray (1.3)
	pɒbəˈgluməʒ	gloom (1.0)	glue (0.5)	twɪˈdeməf	dame (0.3)	day (2.9)
	gadʒɪˈwumək	womb (1.0)	woo (0)	vɪˈlenəθ	lane (1.5)	lay (1.7)
	flɪˈzumɪp	zoom (0)	zoo (1.0)	sɪgləˈstenɪv	stain (0.8)	stay (1.6)
	səˈglimɪb	gleam (0.7)	glee (0.6)	wəˈdoməb	dome (1.0)	doe (0.5)
	bləˈkinɪm	keen (1.4)	key (1.9)	gakəˈfomɪp	foam (0.9)	foe (1.2)
	tɒʒəˈskimɪð	scheme (1.8)	ski (0.7)	dɪˈhomɪg	home (2.7)	hoe (0.3)
		mean 0.97	***mean 1.18***		***mean 1.04***	***mean 1.57***
		(sd 0.75)	***(sd 1.00)***		***(sd 0.82)***	***(sd 0.77)***
Maximum-lengthening contexts (voiced fricatives)	klatɪˈbuzɪv	booze (0.5)	boo (0)	θəˈloðɪp	loathe (0)	low (2.2)
	nʌgəˈtʃuzɪð	choose (1.2)	chew (0)	bɪvɪˈstovəʃ	stove (1.2)	stow (0)
	θapəˈfjuzɪb	fuse (0.7)	few (2.8)	ðaplɪˈwovədʒ	wove (0)	woe (0.5)
	kɪˈgruvɪp	groove (0.6)	grew (1.3)	dʒəˈbeʒɪm	beige (0.5)	bay (1.5)
	fləˈmuvɪdʒ	move (1.9)	moo (0)	vɒglɪˈgevətʃ	gave (1.9)	gay (1.1)
	gɪˈmjuzəʃ	muse (0)	mew (0)	bɪˈpevɪð	pave (0)	pay (1.8)
	pləˈruʒəg	rouge (0)	rue (0)	hɛtʃɪˈfezɪʒ	phase (1.5)	fay/fey (0.5)
	frʌnɪˈsiðɪg	seethe (0)	see (2.5)	faθɪˈrezɪdʒ	raise (1.0)	ray (0.5)
	ʃɪkləˈtiðɪʒ	teethe (0)	tea (2.0)	kwɪˈsevɪg	save (1.3)	say (2.3)
	wəˈlivɪʒ	leave (1.8)	lee/lea (1.3)	ʃɒgəˈslevɪm	slave (1.2)	slay (0)
	nəˈbriðɪv	breathe (0.6)	Brie (0)	sləˈwevɪs	wave (1.6)	way (3.1)
		mean 0.67	***mean 0.89***		***mean 0.93***	***mean 1.22***
		(sd 0.71)	***(sd 1.10)***		***(sd 0.69)***	***(sd 1.03)***