JOURNAL ARTICLE

Prevocalic t-glottaling across word boundaries in Midland American English

Kamil Kaźmierski

Department of Contemporary English Language, Faculty of English at Adam Mickiewicz University, Poznań, PL
kamil.kazmierski@wa.amu.edu.pl

Rates of t-glottaling across word boundaries in both preconsonantal and prevocalic contexts have recently been claimed to be positively correlated with the frequency of occurrence of a given word in preconsonantal contexts (Eddington & Channer, 2010). Words typically followed by consonants have been argued to have their final /t/s glottaled more often than words less frequently followed by consonants. This paper includes a number of 'internal' and 'external' predictors in a mixed-effects logistic regression model and has two goals: (1) to replicate the positive correlation of the frequency of occurrence of a word in preconsonantal contexts (its 'contextual frequency') with its rates of t-glottaling in both preconsonantal and prevocalic contexts postulated by Eddington and Channer (2010), and (2) to quantify the factors influencing the likelihood of t-glottaling across word boundaries in Midland American English. The effect of contextual frequency has been confirmed. This result is argued to support a hybrid view of phonological storage and processing, one including both abstract and exemplar representations. T-glottaling has also been found to be negatively correlated with bigram frequency and speech rate deviation, while positively correlated with young age in female speakers.

Keywords: contextual frequency; t-glottaling; Midland American English

1. Introduction

T-glottaling is the realization of the voiceless alveolar plosive /t/ as a glottal stop [ʔ]. It is also sometimes referred to as 'glottal replacement' (Milroy, Milroy, Hartley, & Walshaw, 1994), to distinguish it from 'glottal reinforcement' (Higginbottom, 1964), that is from the addition of [ʔ] to an oral stop. T-glottaling has been attested in numerous accents of English spoken both in Britain (Fabricius, 2002; Milroy et al., 1994; Trudgill, 1974) and in the United States (Eddington & Channer, 2010; Eddington & Taylor, 2009; Huffman, 2005; Levon, 2006; Roberts, 2006; Seyfarth & Garellek, 2015). Recently, it has been shown that in American English, [ʔ] is one of the possible realizations of /t/ across word boundaries before vowels, that is in contexts such as *righ*[ʔ] *in*. Such realizations are attested by Levon (2006), in the speech of New York Reform Jews, by Roberts (2006), for rural Vermonters, and by Eddington and Taylor (2009), who compared Utahns to speakers from elsewhere in the U.S. Eddington and Channer (2010) argued that the likelihood of t-glottaling is linked to the frequency with which a given /t/-final word occurs before consonant-initial words. This is an intriguing proposition. The evidence that the original paper provides for this proposition, however, can be further strengthened. Using different data, an improved statistical analysis, and a more precise method of calculating the proportion of occurrence of a word in preconsonantal contexts, the present paper confirms the finding of Eddington and Channer (2010): Higher rates of occurrence in preconsonantal contexts

correlate with a higher chance of glottaling in both preconsonantal and prevocalic contexts. Consequently, the case that the original paper made for storage of phonetically detailed representations is strengthened. The contribution of the present study is twofold. First, it investigates the influence of linguistic and social factors on the likelihood of prevocalic t-glottaling across word boundaries in Columbus, Ohio. Second, it bolsters the support given by Eddington and Channer (2010) to models of phonology that incorporate phonetically rich representations by providing further evidence that t-glottaling shows a 'contextual frequency' effect (Forrest, 2017).

1.1. Glottaling

The glottal stop [ʔ], in its canonical realization, is a plosive produced by drawing together, and then releasing, the vocal folds. Glottal stops, however, show considerable variability in their realization, ranging from full stops to laryngealized phonation (Garellek, 2013, p. 2). In fact, the most common realization of [ʔ] in English is that of "irregular spacing of pitch periods in the waveform" (Dilley, Shattuck-Hufnagel, & Ostendorf, 1996, p. 428; Pierrehumbert & Talkin, 1992), a type of 'incomplete [ʔ]' identified by Garellek's (2013, p. 33–54) articulatory study. Consequently, in this study I use the term 'glottal stop' in this broader sense, including both full stops and laryngealized phonation. An example of t-glottaling—intervocalic /t/ at a word boundary realized as a form of [ʔ]— is presented in **Figure 1**. It shows a case of t-glottaling from the Buckeye corpus (Pitt et al., 2007), the source of data for the present study. The waveform and spectrogram show /t/ realized as [ʔ] by a young woman from Columbus, Ohio. A token of [ʔ] is visible as irregular pitch periods. Though the [ʔ] is 'incomplete,' such realizations do give the auditory impression of a glottal stop, not unlike that well-attested for British English accents (Fabricius, personal communication). As for possible sources of this auditory impression, dips in f0 and amplitude have been identified as cues that listeners rely on for detecting glottal stops even when there is no cessation of voicing (Hillenbrand & Houde, 1996).
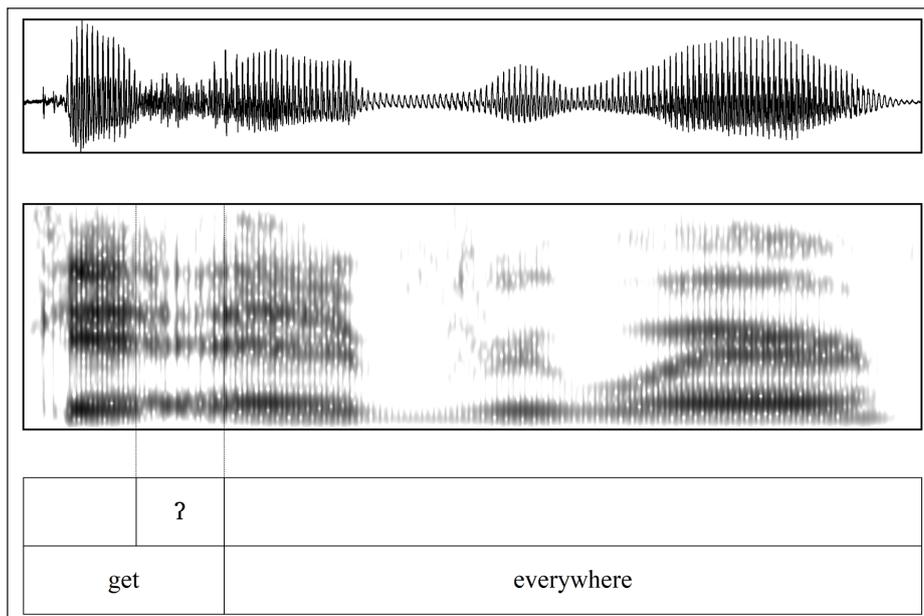

Figure 1: Waveform and spectrogram of *get everywhere* produced by a young woman from Columbus, Ohio (Source: the Buckeye corpus, Pitt et al., 2007). The portion of the signal associated with the percept of [ʔ] shows irregular pitch periods.

T-glottaling has long been observed to take place in American English dialects in several phonological environments (see **Table 1** for a brief overview of reports). As is the case with other consonantal features, t-glottaling has been described more extensively for British English than for American English accents (though the number of studies on variation in consonants in American English accents is growing, e.g., Zhao, 2010; Gylfadottir, 2015; Yuan & Liberman, 2011). T-glottaling has been reported to occur word-internally (e.g., Trager, 1942; Zue & Laferriere, 1979), across word boundaries before sonorant consonants (Pierrehumbert, 1994) and before plosives (Huffman, 2005). It has only relatively recently been reported to occur across word boundaries before vowels, as in *righ*[ʔ] *ankle* (Eddington & Taylor, 2009; Levon, 2006; Roberts, 2006).

As Eddington and Channer (2010) observe, prevocalic glottaling across words may be seen as surprising in American English, when one considers the rarity of prevocalic glottaling word-internally. Within words, prevocalic glottaling is generally less commonly reported for North American than for British varieties: In contexts such as *ci*[ɾ]*y*, flapping typically takes place, precluding glottaling. Although word-internal prevocalic t-glottaling is attested in American dialects—it has been observed before /ən/ in dialects which 'unpack' syllabic /n̩/, as in *moun*[ʔə]*n* (Eddington & Savage, 2012), and also, to a limited extent, in morphologically complex words such as *pu*[ʔ]*ing, wai*[ʔ]*ing* (Patterson & Connine, 2001)—word-internal prevocalic flapping is more common than word-internal prevocalic glottaling. In general, glottaling and flapping tend to occur in mutually exclusive contexts, as elaborated in the following. Some environments allow only glottaling. This is the case word-internally before obstruents and nasals, as in *ou*[ʔ] *put* or *po*[ʔ]*ent*, as well as in absolute final position, e.g., in *tha*[ʔ]. Before vowels and syllabic liquids within words, on the other hand, flapping is typical, as in *ci*[ɾ]*y, be*[ɾ] *er*, and *li*[ɾ]*le*.[1] Across word boundaries, only glottaling is allowed before obstruents, as in *se*[ʔ] *back*, and before sonorant consonants, as in *abou*[ʔ] *you*. But across word boundaries before vowels—the environment investigated in the present paper—both flapping and glottaling are common: *righ*[ɾ] *around* ~ *righ*[ʔ] *around* (Eddington & Channer, 2010).

Eddington and Channer (2010) present a large-scale study of the competition of glottaling and flapping in unscripted speech. They analyzed a sub-part of the Santa Barbara Corpus of Spoken American English (DuBois et al., 2005), looking at all cases of word-final /t/ (preceded by a vowel, nasal, or liquid) followed by vowel-initial words ($N = 1,101$). In that study, all instances of /t/ were impressionistically coded as [t], [ɾ], [ʔ], or as deleted.

Table 1: Overview of contexts for which t-glottaling has been reported to take place in American English.

Context	Examples	Sources
before nasals	*por*t*ent*	Trager (1942), Zue and Laferriere (1979)
before sonorants	*ha*t *rack, abou*t *you*	Pierrehumbert (1994), Kaźmierski et al. (2016)
before plosives	*bee*t *counter*	Huffman (2005), Dilley and Pitt (2007)
syllable-finally	*ca*t*, sen*t*, bel*t	Selkirk (1972), Kahn (1976), Cohn (1993)
IP-finally	*'ha*t*, daypack'*	Huffman (2005)
across words, before vowels	*righ*t *around*	Levon (2006), Eddington and Taylor (2009)

---

[1] Trager (1942) reports having [ʔ] in *prattle* and *glottal*, but not in *little* and *bottle*.

The middle column of **Table 2** shows the results. In a non-negligible number of cases (262 out of 1,101, that is just under 24%), prevocalic /t/ was realized as a glottal stop. This might be seen as surprising, given the prevalence of prevocalic flapping word-internally in this segmental context.

1.2. Contextual frequency effects

Eddington and Channer (2010) posit that the /t/ which glottalizes before vowels behaves in a way mimicking the preconsonantal position: T-glottaling is very widespread before consonants, both word-internally and across word boundaries. Words ending in /t/ which are typically followed by consonant-initial words, they argue, store numerous exemplars with [ʔ], and these exemplars influence production in that they raise the likelihood of a [ʔ] realization regardless of context. As a result, the overall likelihood of t-glottaling in these words increases, even when occurring before vowels. Thus, the effect of the frequency of occurrence in preconsonantal environment on t-glottaling can be seen as a case of a 'contextual frequency' effect (cf. Forrest, 2017). There is indeed a growing body of evidence supporting the hypothesis that the mental representations of lexical items include fine phonetic detail, whose shape is driven by the phonological environment in which lexical items typically occur. Such effects, also known as 'cumulative context effects' (Raymond, Brown, & Healy, 2016) have been hypothesized to be influenced by 'Frequency in Favoring Conditioning' (Bybee, 2017) or 'Frequency in a Favorable Context' (Brown & Raymond, 2012); phonetic shapes of lexical items are influenced by the frequency of occurrence in contexts which favor a particular realization, rather than by overall lexical frequency. These findings challenge long-accepted views of lexical storage which assume that all non-contrastive information is abstracted away. Such 'abstract only' models of speech production assume a feed-forward, modular architecture. In phonological research, this view of phonological storage has been standard since at least SPE (Chomsky & Halle, 1968), and in further iterations of generative phonology such as Optimality Theory (Prince & Smolensky, 1993). A fully spelled-out speech production model assuming both modularity and abstractness is proposed in Levelt et al. (1999). It posits that the phonological form of a lexeme is constructed from abstract phonemes. This is done after lexical retrieval and morphological operations, so the phonological form cannot be influenced by lexical identity or morphological composition. Any lexical-identity effects that are allowed are word-form level frequency effects. This 'abstract only' view of speech production has faced a number of challenges which call into question the assumptions that the phonological level operates on abstract units devoid of phonetic detail, and that speech production proceeds in discrete modules in a feed-forward fashion. The following findings exemplify these challenges. First, effects of lemma frequency on phonetics such as those shown by Gahl (2008), who found that homophones such as *time* and *thyme* show differing degrees of reduction are problematic for modularity. Under the modularity assumption, once the phonological level has been reached, no semantic

Table 2: The realization of word final /t/ in Eddington and Channer (2010), based on a sub-part of the Santa Barbara Corpus and in the present data set, based on the Buckeye corpus.

Realization of /t/	Eddington and Channer (2010)	Buckeye
Flapping [ɾ]	656 (59.6%)	4,669 (63.8%)
Glottaling [ʔ]	262 (23.8%)	1,134 (15.5%)
No process [t]	110 (10%)	763 (10.4%)
Deletion [∅]	73 (6.6%)	751 (10.3%)

information should be accessible; a particular string of phonemes should be passed on to the phonetic component, and so no difference between *time* and *thyme* is expected. While word-form frequency effects could be accommodated by the feed-forward modular architecture, lexeme frequency effects cannot. Second, there is evidence of morphological information influencing phonetics: e.g., Strycharczuk and Scobbie (2016) found different degrees of GOOSE-fronting in apparent homophones such as *ruler* and *rul+er*. Again, the phonological string is identical in each case, and the influence on phonetics of a module preceding phonology challenges modularity and exclusively abstract storage. Finally, there are the cumulative context effects, where the typical environment in which a lexeme occurs influences its phonetic realization. Seyfarth (2014) found that durations of words which are typically predictable from their immediate context (e.g., *current*) are reduced more than durations of words which are typically less predictable from their immediate context (e.g., *nowadays*), even when a particular instance of the typically predictable word (e.g., *current*) is not, in fact, predictable. Baumann and Ritt (2017) show that the development of the link between morphosyntactic category and word stress in English of the ˈ*research* – *re*ˈ*search* type can be modeled by assuming that lexical stress is an accumulation of repeated adaptations to phrase-level rhythm.

Eddington and Channer's (2010) finding poses another serious challenge to 'abstract only' models. If the likelihood of occurrence of a specific allophonic variant is influenced by the occurrence of a particular lexeme in a particular environment, then the storage of a phonetically rich representation is required. However, the statistical analysis employed in the original study leaves it open to criticism. In the first part of their statistical analysis, the authors fitted a logistic regression model to assess the influence of several variables on the likelihood of t-glottaling: The realization of a final /t/ in a given word was a binary response variable (glottaling versus any other realization). This model, however, did not include a measure of the typical environment of a word among the predictors. The hypothesis that prevocalic t-glottaling across word boundaries is driven by a contextual frequency effect was then tested separately, outside of the regression model, in the second part of the statistical analysis. In it, the authors measured the proportion with which each of the test words was followed by consonant-initial words. To get reliable estimates, they used a corpus much larger than the Santa Barbara Corpus: the Corpus of Contemporary American English (COCA) (Davies, 2010). Words realized with [ʔ] by the speakers were found to be followed by a consonant on average 64% of the time in the large corpus, whereas words realized with other allophones by the speakers were found to be followed by a consonant on average 60% of the time. Using ANOVA, the authors determined this difference to be statistically significant, and drew the conclusion that there is a relationship between the frequency of occurrence before consonants and the likelihood to undergo glottaling. A serious limitation of this approach is that the second part of the analysis did not control for potential confounds. The authors themselves demonstrate the influence of a number of factors on the likelihood of glottaling with their regression model in the first part of their analysis. These possible confounds, however, are not taken into account in the second part, the ANOVA test. Furthermore, the observations in the data set were not independent as they included both multiple tokens of the same words, and multiple words uttered by the same speakers. These two problems increased the risk that the significant result was a false positive. The present study addresses both these issues. The issue of confounding factors is addressed by including the proportion of consonant-initial words following a given test word in a large corpus (henceforth CONSONANTAL PROPORTION) as one of several predictors in a mixed-effects logistic regression model featuring a number of other predictors, known from prior research to be of relevance for t-glottaling. The

issue of non-independence of observations is addressed by incorporating random terms for participants and words into the model architecture.

2. Data

All bigrams—sequences of one word (w_1) immediately followed by another word (w_2)—in which w_1 ends in /-Vt/ and w_2 begins with a vowel were retrieved from the Buckeye corpus (Pitt et al., 2007) with the help of LaBB-CAT (Fromont & Hay, 2012), a publicly available corpus annotation and management suite. The Buckeye corpus is very well-suited to this study as it contains unscripted speech of a homogenous group of speakers of a variety for which t-glottaling has not been the center of attention. It also comes with hand-corrected phone-level annotations, including allophones of /t/ (an evaluation of the accuracy of the transcriptions is described below). Cases where w_2 was not a lexical item, that is when it was represented as one of the following orthographic forms: *um*, *uh*, *oh*, *o*, *uh-huh* were excluded from analyses. This resulted in 7,317 hits, occurring in 5,375 speaker turns. The Buckeye Speech Corpus[2] contains recordings of 40 speakers from Columbus, Ohio. The participants were recorded conversing with an interviewer on everyday topics. The recordings are altogether around 40 hours long and contain approximately 300,000 word tokens. The city of Columbus, according to the dialectal division of North America of Labov, Ash, and Boberg (2006) belongs to the Midland dialect region. Speech patterns in this region can be expected to be typical of North American English, to the extent that any particular area can. "Many features of the Midland are the default features—that is, the linguistic landscape remaining when marked local dialect features are eroded" (Labov et al., 2006). While there are vocalic features that nonetheless separate the Midland dialect from the neighboring Northern and Southern dialect areas (such as the 'close approximation' of the vowels in the LOT and THOUGHT lexical sets, Labov et al., 2006, p. 264), consonantal features, or glottaling specifically, is not known to show any particular pattern in this dialect. The speakers in the corpus are stratified for age (two categories: under 30 and over 40, actual ages are not provided to the corpus user) and gender.

The corpus comes with, among others, a phone-level annotation layer produced by "a group of trained phonetic annotators […] paid for corpus preparation" (Dilley & Pitt, 2007, p. 2341), based on spectrograms, waveforms, and auditory cues. The phone-level annotation layer includes /t/ allophony. A segment was labeled by the corpus annotators as [ʔ] if it "had perceptually creaky voicing accompanied by irregularity in pitch period timing in the waveform" (Dilley & Pitt, 2007, p. 2342). This study relies on these transcriptions, as there are good reasons to believe that the accuracy of coding of /t/ allophony in Buckeye's segment annotations is high. First, in a published study (Pitt, Johnson, Hume, Kiesling, & Raymond, 2005), the intertranscriber reliability for stops was found to be 93%. Second, in a study investigating sequences of word-final /t/ followed by /j/-initial words (/V_#j/) e.g., *about you* (Kaźmierski, Wojtkowiak, & Baumann, 2016), 65% of /t/s (528/808) were hand-coded as glottaled, based on a drop in amplitude and lack of release burst. A look at the transcriptions provided with the corpus reveals that 54% (446/833) are labeled as [ʔ] (the slight difference in *N* seems to stem from differences in querying with the bundled software versus with LaBB-CAT). Finally, in a study even more directly related to the present one, Seyfarth and Garellek (2015) inspected a subset of coda /t/s in the Buckeye corpus and found that the vast majority of [ʔ] labels (1,762/1,824, or 97%) correspond to glottal replacement, with no format transitions or release bursts indicative of [t]. The remaining small minority of cases (62/1,824, or 3%) correspond to glottal reinforcement

---

[2] Available at http://buckeyecorpus.osu.edu.

[ʔt̪], as they have a detectable release burst.[3] Taken together, these results suggest that the accuracy of transcriptions of [ʔ] is high, and that, if anything, [ʔ] could be under- rather than over-reported in Buckeye's annotations.

All cases where no consonantal sound was present as a transcription of the final sound of w_1 were treated as cases of deletion. Cases where the final consonant of w_1 was transcribed as a voiced plosive /d/ were included together with those transcribed as a flap. The remaining cases, where the final sound of w_1 was transcribed with a plain 't' symbol were coded as [t], denoting a voiceless alveolar plosive. The rightmost column of **Table 2** shows a general overview of the initial data set ($N = 7{,}317$). A comparison of proportions of the realizations of /t/ in Eddington and Channer's (2010) results and in the present data set shows a lower glottaling rate, and higher flapping and deletion rates for Buckeye.

For further analysis, including regression modeling, only cases where /t/ was realized either as a flap or glottal stop were kept ($N = 5{,}803$). This is motivated by the observation that it is these two allophones of /t/ that compete directly in word-final prevocalic position (cf. Eddington & Channer, 2010, p. 344). The realization of /t/ as [ɾ] or [ʔ] was therefore included as a binary response variable GLOTTALED, with [ɾ] as a reference level, so that effectively the likelihood of glottaling over flapping is modeled.

Figure 2 shows the final data set broken down by gender and age. A glance at the raw data shows that there are no speakers categorically preferring only one realization—everyone
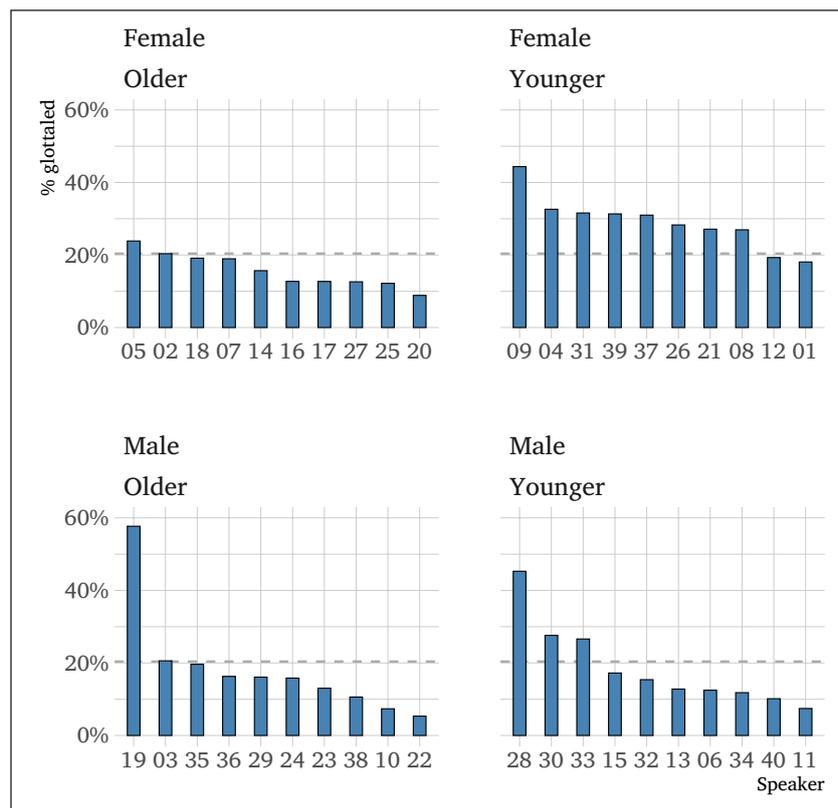

Figure 2: Individual variation broken down by age and gender. Each bar represents the percentage of glottaling for a single speaker after only glottaling and flapping were left in the data set ($N = 5{,}803$). The dashed line represents the mean glottaling rate for all speakers ($\bar{x} = 0.204$). Speaker indices are shown on the horizontal axis.

---

[3] The difference between glottal reinforcement and glottal replacement is not always unambiguous based on acoustics alone, however. The adduction of the vocal folds completely overlapping with the alveolar closure [ʔt̚] would mask the latter. The prevalence of such masked articulations is unknown, though Huffman (1998) found that in word-final prepausal position 70% of /t/s were realized as [ʔt̪], with the remaining 30% split between [t] and [ʔ].

shows variability between the two. Further, younger female speakers seem to have higher rates of glottaling than the three other groups, with eight out of ten younger female speakers having a glottaling rate above the mean. Going against this general trend, Speaker 19, an older male speaker, has the highest glottaling rate in the data set.

3. Analysis

In any corpus of unscripted speech, the data is, by definition, not controlled at the collection stage, allowing confounding variables to be present. Thanks to multiple regression modeling, where several predictors are included in the same model, confounds can be dealt with at the stage of statistical analysis (Baayen, 2008). In other words, the influence of a variable of theoretical interest can be estimated while the influence of the other predictors is kept constant. In the present analysis, control variables were added mostly through automatic annotation functionality of LaBB-CAT (Fromont & Hay, 2012), as well as data transformation functions of the R package DPLYR (Wickham, Francois, Henry, & Müller, 2017), as discussed separately for each variable below. A mixed-effects logistic regression model was then fitted to the data with the *glmer*() function from the LME4 package (Bates, Mächler, Bolker, & Walker, 2015) in R (R Core Team, 2018). The inclusion of random effects, that is the use of mixed-effects modeling, is appropriate where several data points are grouped, as is the case here, and is known to prevent the inflation of the rate of Type I errors (Baayen, Davidson, & Bates, 2008). Accordingly, to account for within-speaker variation that is not due to any of the variables listed in Sections 3.1–3.2 below, and to account for the lack of independence of multiple data points coming from the same speaker, SPEAKER was included as a by-subject random intercept term. This takes care of differences in rates of glottaling across speakers. Furthermore, as the influence of the test variable CONSONANTAL PROPORTION (see Section 3.1 below for details on the test variable) might be different on different speakers, a by-subject random slope for CONSONANTAL PROPORTION was included (while including by-subject random slopes for all fixed terms would be desirable, it resulted in singular fit and was therefore not feasible). Similarly, word-level idiosyncrasies could not be ruled out, and were accounted for by including WORD, that is each /t/-final test word, as a by-item random intercept term. Bringing the influence of confounds under statistical control by including them as fixed effects, as well as accounting for interdependencies between data points by including random intercepts, is seen as a considerable improvement compared to the ANOVA test of the study by Eddington and Channer. Finally, as a remedy to initial model convergence issues, the BOBYQA optimizer was used.

3.1. Test variable: Consonantal proportion

The main variable of interest for this study is the proportion of consonant-initial words following a given word in a corpus. The motivation for investigating its influence on t-glottaling comes from Eddington and Channer (2010), yet the details of how it was computed have been slightly improved compared to that study. Here, it was computed based on the SUBTLEX-US (Brysbaert & New, 2009) corpus, a large collection of movie and TV series subtitles, shown to yield frequency measures that give accurate predictions in reaction time experiments. With the help of AntConc, a freeware corpus analysis toolkit (Anthony, 2014), all bigrams where the test word appeared as the first word were retrieved. Next, all words were supplied with a phonological transcription from the CMU pronouncing dictionary with the LOGIOS Lexicon Tool.[4] Finally, for each test word the number of consonant-initial following words was divided by the sum of consonant-initial

---

[4] Available at http://www.speech.cs.cmu.edu/tools/lextool.html.

following words and vowel-initial following words. By using phonological transcriptions instead of orthography, which was the case in the original study, higher accuracy has been achieved, as words with a consonant letter in spelling where none is present in pronunciation (e.g., *hour*) and words with no consonant letter in spelling where there actually is one in pronunciation (e.g., *use*) were classified correctly.

As is illustrated in **Figure 3**, the /t/-final words analyzed in the present paper tend to be followed by consonant-initial words more often than by vowel-initial words, based on SUBTLEX-US: The mass of the density in both panels is above 0.5. This is the case both in terms of token counts (mean = 0.672, median = 0.669, *SD* = 0.142, *n* = 5,803) and in terms of type counts (mean = 0.685, median = 0.696, *SD* = 0.111, *n* = 209). Incidentally, the assumption that glottaling is favored before consonants is bolstered by the statistics of glottaling of post-vocalic word-final /t/ across word boundaries in the Buckeye corpus. When the following word starts with a consonant, 5,112 out of 14,124 cases (36%) are glottaled, and when the following word starts with a vowel, only 870 out of 7,323 cases (12%) are glottaled.

This continuous variable was centered (by subtracting the mean) and standardized (by dividing by one standard deviation), giving CONSONANTAL PROPORTION present in the tables and charts to follow. The hypothesis with regard to this variable, and so the main hypothesis of the study, is that it will be positively associated with the likelihood of a test word undergoing t-glottaling.

3.2. Control variables

Based on a review of existing relevant research on speech variation generally and t-glottaling specifically, the following control predictors were included in the model.

3.2.1. The frontness of the initial vowel of w_2

There are conflicting findings in the literature with regard to the influence of the frontness of the initial vowel of the following word on the likelihood of glottaling. On the one hand, Eddington and Taylor (2009) suggest that glottaling is favored by following front vowels, while Ostalski (2013) seems to have found the opposite. Whichever way the direction goes, these findings at the very least suggest that the frontness of the initial vowel of w_2 may be of some relevance to glottaling, and so it was included in the model as a binary predictor variable w_2 FRONTNESS. It was retrieved automatically from CMU
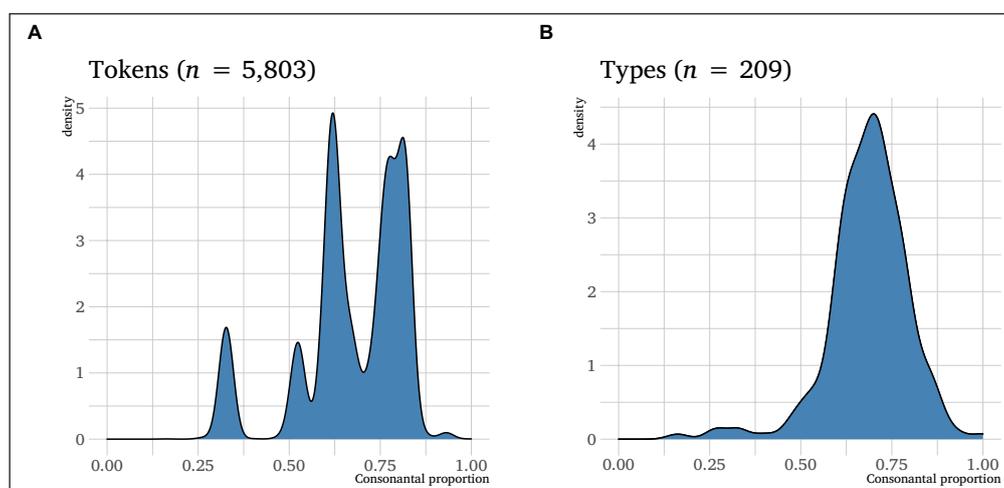

Figure 3: Densities of CONSONANTAL PROPORTION values for words in the analyzed data set (Buckeye, *N* = 5,803), for tokens (left) and types (right). CONSONANTAL PROPORTION values from SUBTLEX-US.

transcriptions. It was entered into the model as a binary, treatment-coded variable (levels: NOT FRONT, FRONT, reference: NOT FRONT).

3.2.2. The presence or lack of stress on the initial syllable of w_2

Previous research suggests that if the word-final /t/ is followed by a stressed syllable, it is more likely to be glottaled (Eddington & Channer, 2010). Therefore, a w_2 STRESS variable was included. It was derived automatically from the CMU transcriptions of each w_2 in the following manner. For monosyllabic words, if it was a function word (either a pronoun, preposition, conjunction, article, or interjection), it was coded as unstressed (unless it was one of the following words: *all*, *each*, *own*). If the monosyllable was a content word, it was coded as stressed. For polysyllables, the initial syllable was coded as stressed if it is transcribed either with primary or secondary stress in CMU. w_2 STRESS was entered into the model as a binary, treatment-coded variable (levels: STRESSED, UNSTRESSED, reference: STRESSED).

3.2.3. Bigram frequency

Words that frequently occur together are more likely to be planned together as a unit compared to word sequences that do not occur frequently next to one another (Bush, 2001; Tanner, Sonderegger, & Wagner, 2017). Being planned as a unit can be seen as making such sequences more word-like, and the phonological behavior at word edges can therefore be expected to approximate word-level phonology. In the present case, if the final /t/ of w_1 in frequent bigrams is subjected to the pressures of within-word phonology, it can be expected to flap more often. To account for this, log-transformed frequencies computed from the results were included as a continuous BIGRAM FREQUENCY fixed effect in the model.

3.2.4. Speech rate

Research on motor planning in speech production suggests that fast speech rate might increase the influence of the neighboring phonological environment on the phonetic shape of a given form (Tanner et al., 2017). The hypothesis behind it is that at higher speech rates, larger chunks of speech are planned together. In the present case, the more likely a given bigram is to be treated as a single unit, the more precedence word-internal restrictions should take. For word-final prevocalic /t/, this suggests that higher speech rates should favor flapping. Speech rate was computed as a number of syllables (taken from CELEX2, Baayen, Piepenbrock, & Gulikers, 1995) in a given speaker turn divided by the length of that turn in seconds, yielding a syllables-per-second speech rate measure. The resulting values were log-transformed and centered separately for each speaker by subtracting from each value a given speaker's mean. As a result, a SPEECH RATE DEVIATION variable was produced, reflecting the hypothesis that it is speeding up or slowing down relative to one's habitual speech rate that might influence rates of t-glottaling, rather than that habitual speech rate itself (cf. Tanner et al., 2017).

3.2.5. Age

There is some indication in the literature that rates of t-glottaling are sensitive to the age of the speaker. In Eddington and Channer (2010), the rates were lower for speakers aged 30 and older. Such age-stratification might indicate change in progress. However, as is always the case with apparent-time data, age grading, namely "a regular change of linguistic behavior with age that repeats in each generation" (Labov, 1994, p. 46) cannot be excluded. Roberts (2006), in her study of a more homogeneous group of speakers—47 Vermonters—reports a more complex pattern. In her study, adolescents and older speakers have higher rates of glottaling compared to the middle group, namely parents. This could suggest age grading of glottaling in this community, as the group most likely to conform

with societal pressures seems to avoid the (locally) stigmatized variable (Roberts, 2006, p. 231). Not having a hypothesis as to how age of the speakers from Columbus might influence their glottaling rate, AGE was still included as a binary treatment-coded variable (levels: YOUNGER [<30] and OLDER [>40], reference: OLDER, according to the speaker metadata present in the corpus) to account for the possibility that it does play a role.

3.2.6. Gender
In an early study of glottaling in the United States, Byrd (1994) found that the overall use of glottal stops was greater for women than for men, regardless of position in a word. On the other hand, glottals were more frequent among male than female speakers in the studies of Levon (2006) and Roberts (2006). This difference could reflect the social salience of glottaling in the communities studied there, as socially salient variables often show gender stratification (Trudgill, 1974; Labov, 1990). For the present data set, as no research is available as to the influence of gender on glottaling in Midland American English, there is no specific hypothesis about this relationship. However, given its role in language variation in general, and its influence on t-glottaling in other communities specifically, GENDER was included as a binary treatment-coded variable (levels: FEMALE, MALE, reference: FEMALE).

3.2.7. Interaction of age and gender
On top of having age and gender as fixed effects, there is also a good reason to include an AGE:GENDER interaction term in the model. Eddington and Taylor (2009) found that younger female speakers were the gender/age combination that glottaled the most of the four possible combinations. As **Figure 2** suggests, the present data set might point in the same direction. Therefore, the inclusion of this interaction term is seen as theoretically justified. This finding, incidentally, might be seen as an indication of an ongoing sound change, as young women tend to be the leaders of change (Eckert, 1988; Labov, 1994).

4. Results
The results of regression modeling are summarized in **Table 3**. As is standard practice, *p*-values for this model were calculated based on asymptotic Wald tests. Each coefficient represents the estimated change in log-odds of glottaling over flapping as the value of the predictor increases (for continuous predictors) or when the level changes from reference to the one indicated in brackets (for categorical predictors). For AGE [YOUNGER], the coefficient obtains when GENDER is held at FEMALE, and for GENDER [MALE] the

Table 3: Coefficient table of the mixed-effects logistic regression model of the probability of glottaling.

Term	Estimate	*SE*	*z* value	*p*-value
(Intercept)	−1.65	0.24	−6.84	<0.001
CONSONANTAL PROPORTION	0.43	0.14	3.08	0.002
W2 FRONTNESS [FRONT]	−0.01	0.08	−0.17	0.866
W2 STRESS [UNSTRESSED]	0.40	0.11	3.68	<0.001
BIGRAM FREQUENCY	−0.13	0.03	−4.76	<0.001
SPEECH RATE DEVIATION	−0.12	0.04	−2.79	0.005
AGE [YOUNGER]	0.89	0.28	3.15	0.002
GENDER [MALE]	0.12	0.30	0.39	0.695
GENDER [MALE]: AGE [YOUNGER]	−0.85	0.40	−2.13	0.034

coefficient obtains when AGE is held at OLDER As the table shows, the contribution of a number of predictors has turned out to be statistically significant. Partial-effect plots of all statistically significant terms are presented in **Figure 4**.

4.1. The test variable

Crucially, the variable of prime interest for this study, that is CONSONANTAL PROPORTION, has a positive coefficient ($\beta = 0.43$), with an associated p-value of 0.002. As such, the test variable has been shown to be positively correlated with the likelihood of glottaling over flapping, at a statistically significant level. This effect is visualized in **Figure 4**, Panel A. The influence of contextual frequency, specifically of the frequency of occurrence in preconsonantal environment has therefore been confirmed, replicating the effect found by Eddington and Channer (2010). In the present study, this effect transpired even when other factors known to influence rates of t-glottaling have been accounted for by including them as covariates in the model, and when random effects were part of the model architecture, thus making a stronger case for the contextual frequency effect than the original study did.

4.2. Other linguistic variables

The frontness of the initial vowel of the following word has not turned out to be statistically significant ($\beta = -0.01$, $p = 0.866$), thus supporting neither Eddington and Taylor (2009) nor Ostalski (2013). There are two possible interpretations of this negative result: Either the influence of the frontness of the following vowel is indeed negligible, with previous findings perhaps showing experimental artifacts, or the effect is too small to be detected with the present data set. As vowel quality is not of primary concern here, it will not be pursued further.

With regard to the influence of lexical stress (Panel D in **Figure 4**) t-glottaling has turned out to be more likely if the following syllable is unstressed ($\beta = 0.4$, $p < 0.001$). This result is opposite to that of Eddington and Channer (2010). One possible reason for
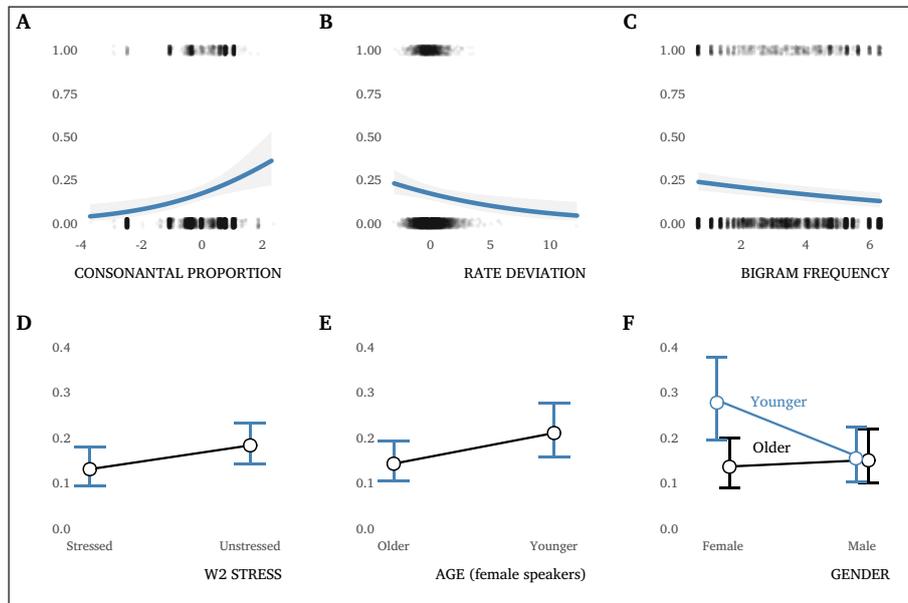

Figure 4: Effect plots for statistically significant model terms. For each term, the effect of the remaining terms was averaged using the EFFECTS package (Fox & Hong, 2009). The y axis in each plot represents the probability of glottaling. Continuous predictors (Panels **A**, **B**, and **C**) include jittered raw data points, included at the bottom (no glottaling) and top (glottaling). To allow the meaningful placement of the raw data points, the full range of probabilities is used in these panels. For categorical predictors (Panels **D**, **E**, and **F**), a selected portion of the probability scale (0–0.4) is plotted for greater readability.

this differing result, suggested by an anonymous reviewer, is the difference in how stress was coded. Eddington and Channer (2010) took phrasal prominence into account when coding stress: For example, function words which "were given an emphatic rendering by the participant" (Eddington & Channer, 2010: 342) were coded as stressed. In contrast, in the present study only lexical stress was considered, and all function words were coded as unstressed (except *all*, *each*, and *own*). Perhaps more importantly, however, an after the fact investigation revealed that the effect of stress in the present study might be confounded by cases where the initial syllable of the following word was realized as a syllabic nasal: 328 such realizations were discovered in the data set. For instance, while the /t/ at the end of *not* in a bigram in the data set, *not intential*, is intervocalic as far CMU dictionary transcription used for retrieving tokens is concerned, its actual realization in the corpus was [nɑʔn̩tɛnʃn̩ʌl]. Syllabic /n/ is known to favor glottaling (Zue & Laferriere, 1979). Indeed, in the present data set, the glottaling rate in cases where the initial segment of w_2 is transcribed as a nasal is 94% (307/328). By comparison, the glottaling rate in the remaining subset of the data is only 15% (827/5,475). In a model fit to the subset of the data set after removing the 328 cases where [ʔ] was not actually intervocalic, the effect of w_2 STRESS went in the same direction as in the original model, but was not significant ($\beta = 0.17$, $p = 0.14$). In the full data set, the glottaling rate is slightly higher before unstressed syllables (936/4,702; 20%) than before stressed syllables (198/1,101; 18%). In the data set with the following nasals removed, however, the glottaling rate is slightly lower before unstressed syllables (646/4,394; 15%) than before stressed syllables (181/1,081; 17%). Note that remaining effects are not substantially different compared to the original model: Crucially, the effect of CONSONANTAL PROPORTION remains almost unchanged at $\beta = 0.41$ ($p = 0.003$). The extent to which syllabic nasals figure into the analysis of the Santa Barbara Corpus in Eddington and Channer (2010) is unknown. At any rate, the effect of following stressed syllables favoring glottaling found by Eddington and Channer, who took phrasal prominence into account when coding for stress, might support the influence of the presence of prosodic boundaries on glottaling. The effect of stress found in the present paper, due to the unfortunate inclusion of a confound, must remain inconclusive.

The corpus frequency of the bigram (BIGRAM FREQUENCY) and SPEECH RATE DEVIATION, had an effect agreeing with that suggested by previous literature. Both are negatively correlated with the likelihood of glottaling: BIGRAM FREQUENCY ($\beta = -0.13$, $p < 0.001$) and SPEECH RATE DEVIATION ($\beta = -0.12$, $p = 0.005$). Seemingly, then, both an increase in speech rate and high bigram frequency are conducive to w_1 and w_2 being chunked together during motor planning, in which case a more word-like behavior of the two-word sequence transpires. In the present case, this means a higher chance of flapping, the process expected word-internally, to occur (cf. Kilbourn-Ceron, Clayards, & Wagner, 2020).

Given the effect of speech rate, as suggested by an anonymous reviewer, one may wonder whether the lower incidence of glottaling in the Buckeye corpus (15.5%) compared to the Santa Barbara corpus (23.8%) may be due to different overall speech rates in the two corpora. To investigate this possibility, I calculated articulation rate, that is the number of syllables per second of phonation, for all speakers in the two corpora. The extraction was automated with the SYLLABLE_NUCLEI script (de Jong & Wempe, 2009), which detects syllable nuclei, discards pauses, and calculates speech rate metrics. The speakers in the Buckeye corpus tend to speak faster (mean = 4.21, median = 4.16, *SD* = 0.425, unit: syllables per second) than speakers in the Santa Barbara Corpus (mean = 3.53, median = 3.81, *SD* = 1.02). As faster speech rate disfavors glottaling, the faster speech rate in the Buckeye corpus may be linked to lower glottaling rates. However, the difference in articulation rates as measured by the SYLLABLE_NUCLEI script might be exaggerated. While

the exact same methodology was applied to both corpora, the script reported suspiciously low articulation rates for some speakers in the Santa Barbara corpus (all speaker means are presented in **Figure 5**). When extreme observations (absolute difference from the mean greater than two standard deviations) are removed, however, the difference in articulation rate between the two corpora persists (Buckeye: mean = 4.19, SD = 0.4, Santa Barbara: mean = 2.72, SD = 0.78).

4.3. Social variables

Predictors relating to social variables were AGE, GENDER, as well as the interaction term GENDER:AGE. The effect of age for the female speakers is statistically significant (β = 0.89, p = 0.002), with the female speakers in the younger age group (<30) showing higher predicted rates of intervocalic t-glottaling than female speakers in the older group (>40) (Panel E in **Figure 4**). The effect of gender for the older age group is not statistically significant (β = 0.12, p = 0.695). However, as the significance of the interaction term between gender and age shows (β = –0.85, p = 0.034), gender does play a role in influencing t-glottaling in that there is an appreciable gender divide among younger speakers (this effect is illustrated in Panel F of **Figure 4**): Younger male speakers glottalize significantly less than younger female speakers. Of the four gender by age group combinations coded in the data set, younger women have the greatest predicted probabilities of t-glottaling. This effect is similar to Eddington and Taylor's (2009) results, but not Eddington and Channer's (2010) results, who found no effect of gender at all. The leading role of young women in t-glottaling in the present study stands in contrast to both Roberts (2006) and Levon (2006), suggesting that the social patterning of t-glottaling for
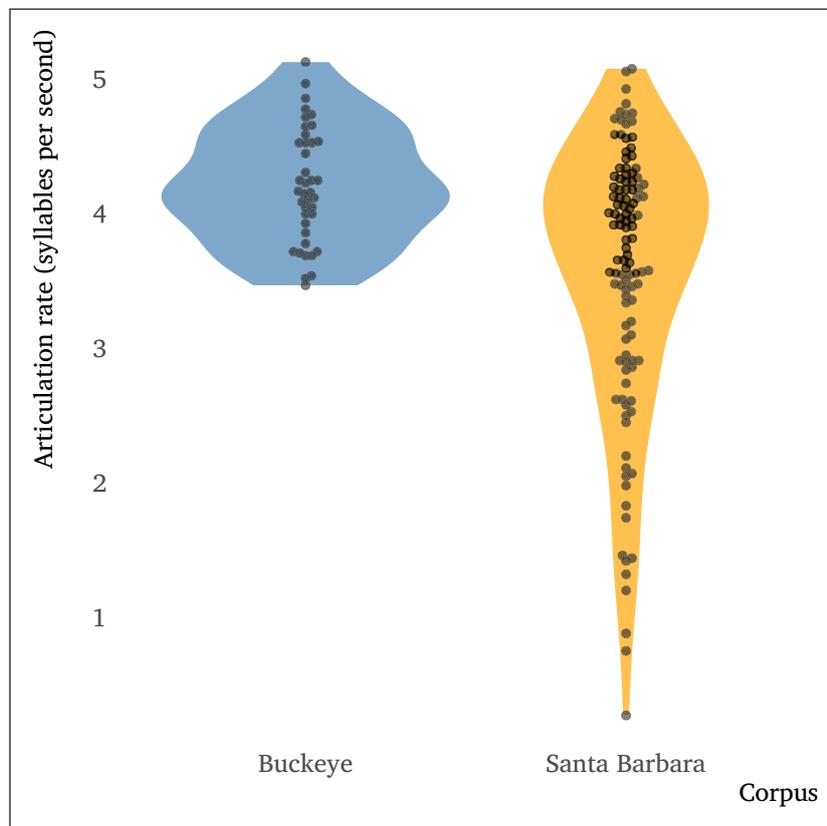

Figure 5: Articulation rates (number of syllables per second of phonation) in the Buckeye Corpus compared to the Santa Barbara corpus extracted with the SYLLABLE_NUCLEI script. Dots are means of individual speakers.

Vermonters and for New York's Reform Jews investigated in the two respective studies is different than for the speakers from Columbus.

5. Discussion and conclusion

With regard to the first goal of the paper, the replication of the contextual frequency effect, it has been confirmed that /t/-final words that typically occur before consonant-initial words undergo glottaling at higher rates than words that occur before consonant-initial words less often. As to the second goal of the paper, the quantification of the factors influencing rates of prevocalic t-glottaling across word boundaries in Midland American English, bigram frequency and speech rate deviation have been found to be negatively correlated with t-glottaling, while the results concerning the frontness of the initial vowel of the following word, and the presence or lack of stress on the initial syllable of the following word remain inconclusive. Finally, an effect of an age by gender interaction was discovered. The implications of the effect of social variables will first be discussed, before turning to the wider implications of the replication of the finding of the contextual frequency effect for models of phonological representations.

In the present study, younger age is positively correlated with t-glottaling for female speakers. While the effect of gender for older speakers is not significant, in light of the interaction that gender enters into with age, it (gender) is by no means an irrelevant variable. While for older speakers there is hardly any difference in rates of intervocalic word-boundary t-glottaling between women and men, for younger speakers there is an appreciable difference between the genders. In Columbus, young women are the group glottaling the most. Variants favored by young women have been repeatedly shown to be the variants on the rise in cases of linguistic change in progress (Eckert, 1988; Labov, 1994). To the extent that this pattern is attested in the present data set, it provides an indication that t-glottaling might be undergoing a change in progress in Columbus. This apparent-time indication, however, would of course have to be supplemented with real-time data to ascertain whether or not t-glottaling is on the rise. The case for a change in progress is strengthened by recurring indications of higher rates of t-glottaling among younger speakers in other parts of the country. Eddington and Taylor's (2009) participants came from Western states ($N = 42$) (the majority of these from Utah) and from non-Western states ($N = 16$, including the North, the South, the Midland, and the Mid-Atlantic), and showed the same kind of gender by age interaction as found here, with younger women having highest rates of t-glottaling. The Santa Barbara Corpus speakers selected by Eddington and Channer (2010) also came from both Western ($N = 19$) and non-Western ($N = 21$) states, and also showed a facilitative effect of age, though not interacting with gender. Studies focusing on particular speech communities provide an interesting comparison. For New York Reform Jews, t-glottaling seems to be tied to their highly specific 'mosaic identity and style' (Levon, 2006), where being affiliated with multiple social groups may exert conflicting pressures: The secular setting of an interview, as well as secular topics, seem to favor glottaling over audible alveolar release for the two young speakers, in contrast to the religious setting of the Youth Group, and religious topics, which disfavor glottaling. Another particularly interesting case is Vermont (Roberts, 2006), where t-glottaling used to be a stigmatized feature, and yet is now spreading among young speakers who otherwise move away from local speech patterns and accommodate to supralocal norms, leading Roberts to posit that the 'new' glottaling might be a new and different feature altogether. All these findings taken together potentially point to a 'nationwide change in progress,' which Midland American English is participating in. Several such supralocal changes have been posited in the vocalic domain. One of them is the low-back (or LOT-THOUGHT) merger, which, at least in the Midland and the South "is not spreading from

any particular point(s) of origin but is appearing roughly simultaneously in several states" (Johnson, 2010, p. 10). Similarly, the Elsewhere Shift (sometimes seen as related to the LOT-THOUGHT merger, cf. Becker, 2019), with the retraction/lowering of the vowels of KIT and DRESS, and with the nasal system of TRAP, recently documented in Lansing, MI (Nesbitt, 2018), was previously found in places so distant from one another as California (Hagiwara, 1997) and Canada (Boberg, 2005).

The hypothesis put forward by Eddington and Channer (2010) has stood its ground when confronted with a different data set and with an improved statistical analysis. The finding that the frequency with which a /t/-final word is followed by consonant-initial words increases the likelihood of that word to undergo glottaling even before vowel-initial words has therefore gained further support. This result contributes to the growing body of evidence that detailed phonetic information is stored in the mental lexicon. The frequency of occurrence in preconsonantal environment can raise the likelihood of glottaling in prevocalic position only if a representation with a glottal stop is stored, as it is not derived online prevocalically. One solution proposed to account for such effects is offered by exemplar models of phonology (e.g., Johnson, 1997; Bybee, 2001). These assume that a phonetically detailed trace of each perceived token is stored.

However, doing away with abstract representations altogether would be a problematic solution: There is after all a vast amount of linguistic data that has led to the development of abstractionist theories to begin with. The pioneering findings of historical linguistics would not have been possible without the stability of lexical classes (as noted for example by Bermúdez-Otero, 2015). Sound changes such as Grimm's Law operated on sound categories shared by multiple lexical items, rather than, or at least in addition to, individual words. Pronunciation of neologisms and loan word assimilations, where sounds of the native language are employed, also attest to the existence of a level of abstract representation (Pierrehumbert, 2001). Consequently, recognizing the need for both abstract and phonetically rich representations, several descriptions of a viable compromise—a hybrid view of representation—have been proposed (Ernestus, 2014; Goldinger, 2007; Nguyen, 2012; Pierrehumbert, 2002, 2006, 2016).

Indeed, research on the perception of word-final /t/ allophony points to the role of both abstract (Sumner & Samuel, 2005) and phonetically rich (Garellek, 2011) representations. Investigating the variation in word-final /t/, Sumner and Samuel (2005) found that all three variants of final /t/ they analyzed, namely [t], [ʔt˺], and [ʔ], cause semantic priming (e.g., [fluːt], [fluːʔt˺], and [fluːʔ] all activate /fluːt/ *flute* and prime *music*). If the non-canonical [ʔt˺] and [ʔ] were not stored, one would expect an advantage of the canonical form [t], as processing the other two variants would involve additional computation or relying on non-canonical cues. If they were stored, one would expect [ʔt˺], the most frequent variant of /t/ prepausally to show an advantage. Neither effect was found: All three variants did better than a mismatched form (e.g., [fluːs]). This null effect is not conclusive with regard to storage. The absence of evidence of a difference cannot be taken as evidence of absence of such a difference: It might be too small to detect with the semantic priming paradigm used by Sumner and Samuel (2005). Indeed, using a phoneme monitoring task, Garellek (2011) found that [ʔt] tokens were recognized faster than canonical [t], suggesting that a more frequent variant of final /t/ is recognized more easily. Sumner and Samuel (2005) further found that only the canonical form, that is [t], shows a long-term priming effect (e.g., [fluːt], but not [fluːʔt˺] or [fluːʔ], primes [fluːt] in a lexical-decision and in a new-old recognition task). This effect, present across two different tasks, supports the primary role of the canonical form, that is of a form with only and all the contrastive features of English phonology, in long-term priming. While this finding of Sumner and Samuel points to the role of abstract representations of final /t/ in perception, the effect found by

Eddington and Channer (2010) and confirmed here bolsters the support for phonetically-rich representations of the same final /t/ in speech production.

Taken together, these two effects support a hybrid model of phonological storage, with both abstract and phonetically rich levels of representation, relevant for different types of processing. To take a brief look at the literature on speech recognition, the role of two levels was demonstrated by several studies (McLennan, Luce, & Charles-Luce, 2003; Vitevitch & Luce, 1998).[5] Tasks involving strong lexical competition (using words with high neighborhood densities, or hard lexical-decision tasks), phonologically ambiguous input, and ample time allowed for recognition, seem to be tapping the abstract level. Other tasks, such as auditory naming, easy-discrimination lexical decision, little time allowed for recognition, seem to be tapping the phonetically rich level. Several hybrid models of phonological storage, promising to accommodate such findings, have been proposed (cf. Ernestus, 2014; Goldinger, 2007; Nguyen, 2012; Pierrehumbert, 2002, 2006, 2016). For the effect confirmed in the present study, detailed representations with a glottal stop and a flap need to be postulated. Interestingly, since in connected speech both of these phonetically quite different variants can occur in the same context, it is clear that their abstract representation with /t/ cannot explain their usage patterns. Instead, phonetically detailed representations become more entrenched with use. This interpretation is consistent with the finding of Dilley and Pitt (2007) that high-frequency /t/-final words show more variability in the realization of /t/ than low-frequency words. Higher lexical frequency entails frequent occurrence in a variety of phonological contexts, and different contexts favor different allophones. The more often a particular word occurs in environments favoring one of the representations, the more entrenched, and the more likely to be used that phonetic form becomes. Hence, a word typically occurring before consonants shows higher glottaling rates, even in prevocalic position. But for the entrenchment in preconsonantal position to take place, the production of [ʔ] must be linked to an activation of an abstract category /t/. Repeatedly, the phoneme finds itself in preconsonantal context, which favors glottaling. Such occurrences strengthen the links between the phoneme in this word, and its allophone [ʔ]. The allophone is then activated even in environments which do not favor this allophone. The stronger the link, the higher the likelihood of glottaling. A functionally-oriented take on t-glottaling, namely that the glottal constriction enhances the recognition of /t/-final words by strengthening the cues to the voicelessness of the coda, was not substantiated (Chong & Garellek, 2018; cf. Seyfarth & Garellek, 2015). Instead, Chong and Garellek (2018) found that glottaling simply does not inhibit the recognition of /t/-final words, while it does inhibit the recognition of /d/-final words. Glottaling, it seems, can spread through the /t/-final words as it does not incur any cost here. These, and other factors influencing t-glottaling can be seen as forming part of the nexus of selection forces on the spread of t-glottaling, one of them being contextual frequency. And, to conclude, while prevocalic t-glottaling presents another case where the 'abstract only' view of phonological representation is insufficient (cf. Baayen, 2007), a hybrid model, rather than an exemplar model of phonological storage is best equipped to accommodate the finding that the rate of glottaling of prevocalic /t/s at word boundaries increases with the frequency of occurrence of a given word before consonant-initial words.

Acknowledgements

---

[5] The differential effect of neighborhood density and phonotactic probability found by Vitevitch and Luce (1998) in speech recognition has also been shown to obtain in production (Kaźmierski, 2019).

References

Anthony, L. (2014). *AntConc [Computer Software]*. Tokyo: Waseda University. Retrieved from http://www.laurenceahttpnthony.net/software

Baayen, R. H. (2007). Storage and computation in the mental lexicon. In G. Jarema & G. Libben (Eds.), *The mental lexicon: Core perspectives* (pp. 81–104). Oxford: Elsevier. DOI: https://doi.org/10.1163/9780080548692_006

Baayen, R. H. (2008). *Analyzing linguistic data*. Cambridge: Cambridge University Press. DOI: https://doi.org/10.1017/CBO9780511801686

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*(4), 390–412. DOI: https://doi.org/10.1016/j.jml.2007.12.005

Baayen, R. H., Piepenbrock, R., & **Gulikers, L.** (1995). The CELEX lexical database (CD-ROM). Philadelphia.

Bates, D., Mächler, M., Bolker, B., & **Walker, S.** (2015). Fitting linear mixed-effects models using *lme4*. *Journal of Statistical Software, 67*(1), 1–48. DOI: https://doi.org/10.18637/jss.v067.i01

Baumann, A., & **Ritt, N.** (2017). On the replicator dynamics of lexical stress: Accounting for stress-pattern diversity in terms of evolutionary game theory. *Phonology, 34*(03), 439–471. DOI: https://doi.org/10.1017/S0952675717000240

Becker, K. (2019). The low-back-merger shift: Uniting the Canadian vowel shift, the California vowel shift, and short front vowel shifts across North America. DOI: https://doi.org/10.1215/00031283-8033373

Bermúdez-Otero, R. (2015). Amphichronic explanation and the life cycle of phonological processes. In P. Honeybone & J. Salmons (Eds.), *The Oxford handbook of historical phonology* (pp. 374–399). Oxford: Oxford University Press.

Boberg, C. (2005). The Canadian shift in Montreal. *Language Variation and Change, 17*(02). DOI: https://doi.org/10.1017/S0954394505050064

Brown, E. L., & **Raymond, W. D.** (2012). How discourse context shapes the lexicon: Explaining the distribution of Spanish f-/h- words. *Diachronica, 29*(2), 139–161. DOI: https://doi.org/10.1075/dia.29.2.02bro

Brysbaert, M., & **New, B.** (2009). Moving beyond Kučera and Francis: A critical evaluation of current word frequency norms and the introduction of a new and improved word frequency measure for American English. *Behavior Research Methods, 41*(4), 977–990. DOI: https://doi.org/10.3758/BRM.41.4.977

Bush, N. (2001). Frequency effects and word-boundary palatalization in English. In J. L. Bybee & P. Hopper (Eds.), *Frequency and the emergence of linguistic structure* (pp. 255–280). John Benjamins. DOI: https://doi.org/10.1075/tsl.45.14bus

Bybee, J. (2001). *Phonology and language use.* Cambridge: Cambridge University Press. DOI: https://doi.org/10.1017/CBO9780511612886

Bybee, J. (2017). Grammatical and lexical factors in sound change: A usage-based approach. *Language Variation and Change, 29*(03), 273–300. DOI: https://doi.org/10.1017/S0954394517000199

Byrd, D. (1994). Relations of sex and dialect to reduction. *Speech Communication, 15*(1–2), 39–54. DOI: https://doi.org/10.1016/0167-6393(94)90039-6

Chomsky, N., & **Halle, M.** (1968). *The sound pattern of English.* New York: Harper & Row.

Chong, A. J., & **Garellek, M.** (2018). Online perception of glottalized coda stops in American English. *Laboratory Phonology: Journal of the Association for Laboratory Phonology, 9*(1), 4. DOI: https://doi.org/10.5334/labphon.70

Cohn, A. C. (1993). Nasalization in English: Phonology or phonetics? *Phonology, 10*, 43–81. DOI: https://doi.org/10.1017/S0952675700001731

Davies, M. (2010). *Corpus of Contemporary American English (COCA): 400+ million words.*

de Jong, N. H., & **Wempe, T.** (2009). Praat script to detect syllable nuclei and measure speech rate automatically. *Behavior Research Methods, 41*(2), 385–390. DOI: https://doi.org/10.3758/BRM.41.2.385

Dilley, L. C., & **Pitt, M. A.** (2007). A study of regressive place assimilation in spontaneous speech and its implications for spoken word recognition. *The Journal of the Acoustical Society of America, 122*(4), 2340–2353. DOI: https://doi.org/10.1121/1.2772226

Dilley, L., Shattuck-Hufnagel, S., & **Ostendorf, M.** (1996). Glottalization of word-initial vowels as a function of prosodic structure. *Journal of Phonetics, 24*(4), 423–444. DOI: https://doi.org/10.1006/jpho.1996.0023

DuBois, J. W., Chafe, W. L., Meyer, C., Thompson, S. A., Englebretson, R., & **Martey, N.** (2005). *Santa Barbara corpus of spoken American English, Parts 1–4.* Philadelphia: Linguistic Data Consortium.

Eckert, P. (1988). Adolescent social structure and the spread of linguistic change. *Language in Society, 17*(2), 183–207. DOI: https://doi.org/10.1017/S0047404500012756

Eddington, D., & **Channer, C.** (2010). American English has go? a lo? of glottal stops: Social diffusion and linguistic motivation. *American Speech, 85*(3), 338–351. DOI: https://doi.org/10.1215/00031283-2010-019

Eddington, D., & **Savage, M.** (2012). Where are the moun[?]ns in Utah? *American Speech, 87*(3), 336–349. DOI: https://doi.org/10.1215/00031283-1958345

Eddington, D., & **Taylor, M.** (2009). T-glottalization in American English. *American Speech, 84*(3), 298–314. DOI: https://doi.org/10.1215/00031283-2009-023

Ernestus, M. (2014). Acoustic reduction and the roles of abstractions and exemplars in speech processing. *Lingua, 142*, 27–41. DOI: https://doi.org/10.1016/j.lingua.2012.12.006

Fabricius, A. (2002). *Ongoing change in modern RP: Evidence for the disappearing stigma of t-glottalling, 23*, 115–136. DOI: https://doi.org/10.1075/eww.23.1.06fab

Forrest, J. (2017). The dynamic interaction between lexical and contextual frequency: A case study of (ING). *Language Variation and Change, 29*(02), 129–156. DOI: https://doi.org/10.1017/S0954394517000072

Fox, J., & **Hong, J.** (2009). Effect displays in R for multinomial and proportional-odds logit models: Extensions to the effects package. *Journal of Statistical Software, 32*(1), 1–24. Retrieved from http://www.jstatsoft.org/v32/i01/. DOI: https://doi.org/10.18637/jss.v032.i01

Fromont, R., & **Hay, J.** (2012). LaBB-CAT: An Annotation store. In *Proceedings of the Australasian language technology association workshop* (pp. 113–117). Dunedin, New Zealand.

Gahl, S. (2008). Time and thyme are not homophones: The effect of lemma frequency on word durations in spontaneous speech. *Language*, *84*(3), 474–496. DOI: https://doi.org/10.1353/lan.0.0035

Garellek, M. (2011). The benefits of vowel laryngealization on the perception of coda stops in English. *UCLA Working Papers in Phonetics*, *109*, 31–39.

Garellek, M. (2013). *Production and perception of glottal stops*. Doctoral Dissertation, Los Angeles: University of California.

Goldinger, S. D. (2007). A complementary-systems approach to abstract and episodic speech perception. In J. Trouvain & W. J. Barry (Eds.), *Proceedings of the 16th international congress of phonetic sciences* (pp. 49–54). Saarbrücken.

Gylfadottir, D. (2015). Shtreets of Philadelphia: An acoustic study of /str/-retraction in a naturalistic speech corpus. *University of Pennsylvania Working Papers in Linguistics*, *21*(2), 89–97.

Hagiwara, R. (1997). Dialect variation and formant frequency: The American English vowels revisited. *The Journal of the Acoustical Society of America*, *102*(1), 655–658. DOI: https://doi.org/10.1121/1.419712

Higginbottom, E. (1964). Glottal reinforcement in English. *Transactions of the Philological Society*, *63*(1), 129–142. DOI: https://doi.org/10.1111/j.1467-968X.1964.tb01010.x

Hillenbrand, J. M., & **Houde, R. A.** (1996). Role of F0 and amplitude in the perception of intervocalic glottal stops. *Journal of Speech, Language, and Hearing Research*, *39*(6), 1182–1190. DOI: https://doi.org/10.1044/jshr.3906.1182

Huffman, M. K. (1998). Segmental and prosodic effects on coda glottalization. *The Journal of the Acoustical Society of America*, *104*(3), 1818. DOI: https://doi.org/10.1121/1.423446

Huffman, M. K. (2005). Segmental and prosodic effects on coda glottalization. *Journal of Phonetics*, *33*(3), 335–362. DOI: https://doi.org/10.1016/j.wocn.2005.02.004

Johnson, D. E. (2010). *Stability and change along a dialect boundary: The low vowels of Southeastern New England* (Vol. 95). Durham, NC: Duke University Press.

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.

Kahn, D. (1976). *Syllable-based generalisations in English phonology*. PhD thesis, MIT.

Kaźmierski, K. (2019). Durational variation in Polish fricatives provides evidence for hybrid models of phonology. In S. Calhoun, P. Escudero, M. Tabain & P. Warren (Eds.), *Proceedings of the 19th international congress of phonetic sciences, Melbourne, Australia, 2019* (pp. 1997–2001). Canberra, Australia: Australasian Speech Science; Technology Association Inc.

Kaźmierski, K., Wojtkowiak, E., & **Baumann, A.** (2016). Coalescent assimilation across word boundaries in American English and in Polish English. *Research in Language*, *14*(3), 235–262. DOI: https://doi.org/10.1515/rela-2016-0012

Kilbourn-Ceron, O., Clayards, M., & **Wagner, M.** (2020). Predictability modulates pronunciation variants through speech planning effects: A case study on coronal stop realizations. *Laboratory Phonology: Journal of the Association for Laboratory Phonology*, *11*(1), 5. DOI: https://doi.org/10.5334/labphon.168

Labov, W. (1990). The intersection of sex and social class. *Language Variation and Change*, *2*, 205–254. DOI: https://doi.org/10.1017/S0954394500000338

Labov, W. (1994). *Principles of linguistic change. Vol. 1: Internal factors*. Malden, MA: Wiley-Blackwell.

Labov, W., Ash, S., & **Boberg, C.** (2006). *The Atlas of North American English*. Mouton de Gruyter. DOI: https://doi.org/10.1515/9783110167467

Levelt, W. J. M., Roelofs, A., & **Meyer, A. S.** (1999). A theory of lexical access in speech production. *Behavioral and Brain Sciences, 22*, 1–75. DOI: https://doi.org/10.1017/S0140525X99001776

Levon, E. (2006). Mosaic identity and style: Phonological variation among Reform American Jews. *Journal of Sociolinguistics, 10*(2), 181–204. DOI: https://doi.org/10.1111/j.1360-6441.2006.00324.x

McLennan, C. T., Luce, P. A., & **Charles-Luce, J.** (2003). Representation of lexical form. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 29*(4), 539–553. DOI: https://doi.org/10.1037/0278-7393.29.4.539

Milroy, J., Milroy, L., Hartley, S., & **Walshaw, D.** (1994). Glottal stops and Tyneside glottalization: Competing patterns of variation and change in British English. *Language Variation and Change, 6*(03), 327. DOI: https://doi.org/10.1017/S095439450000171X

Nesbitt, M. (2018). Economic change and the decline of raised TRAP in Lansing, MI. *University of Pennsylvania Working Papers in Linguistics, 24*(2), 67–76.

Nguyen, N. (2012). Representations of speech sound patterns in the speaker's brain: Insights from perception studies. In A. C. Cohn, C. Fougeron & M. K. Huffman (Eds.), *The Oxford handbook of laboratory phonology*. Oxford: Oxford University Press.

Ostalski, P. (2013). Glottal stops in General American (intervocalic environments). In E. Waniek-Klimczak & L. R. Shockey (Eds.), *Teaching and researching English accents in native and non-native speakers* (pp. 241–251). Heidelberg: Springer. DOI: https://doi.org/10.1007/978-3-642-24019-5_18

Patterson, D., & **Connine, C. M.** (2001). Variant frequency in American English flap production. *The Journal of the Acoustical Society of America, 109*(5), 2445–2445. DOI: https://doi.org/10.1121/1.4744661

Pierrehumbert, J. B. (1994). Knowledge of variation. In *Papers from the parasession on variation, 30th meeting of the Chicago Linguistic Society* (pp. 232–256). Chicago: Chicago Linguistic Society.

Pierrehumbert, J. B. (2001). Exemplar dynamics: Word frequency, lenition and contrast. In J. Bybee & P. Hopper (Eds.), *Frequency and the emergence of linguistic structure* (pp. 137–157). Amsterdam/Philadelphia: John Benjamins. DOI: https://doi.org/10.1075/tsl.45.08pie

Pierrehumbert, J. B. (2002). Word-specific phonetics. In C. Gussenhoven & N. Warner (Eds.), *Laboratory phonology VII* (pp. 101–139). Berlin: Mouton de Gruyter.

Pierrehumbert, J. B. (2006). The next toolkit. *Journal of Phonetics, 34*, 516–530. DOI: https://doi.org/10.1016/j.wocn.2006.06.003

Pierrehumbert, J. B. (2016). Phonological representation: Beyond abstract versus episodic. *Annual Review of Linguistics, 2*, 33–52. DOI: https://doi.org/10.1146/annurev-linguistics-030514-125050

Pierrehumbert, J., & **Talkin, D.** (1992). Lenition of /h/ and glottal stop. In Gerard J. Docherty & D. Robert Ladd (Eds.), *Papers in laboratory phonology II* (pp. 90–117). Cambridge: Cambridge University Press. DOI: https://doi.org/10.1017/CBO9780511519918.005

Pitt, M. A., Dilley, L., Johnson, K., Kiesling, S., Raymond, W. D., Hume, E., & **Fosler-Lussier, E.** (2007). Buckeye corpus of conversational speech. Columbus, OH. Retrieved from www.buckeyecorpus.osu.edu

Pitt, M. A., Johnson, K., Hume, E., Kiesling, S., & **Raymond, W.** (2005). The Buckeye corpus of conversational speech: Labeling conventions and a test of transcriber reliability. *Speech Communication, 45*(1), 89–95. DOI: https://doi.org/10.1016/j.specom.2004.09.001

Prince, A., & **Smolensky, P.** (1993). Optimality Theory: Constraint interaction in generative grammar. ROA-537.

R Core Team. (2018). *R: A language and environment for statistical computing.* Vienna, Austria: R Foundation for Statistical Computing. Retrieved from https://www.R-project.org/

Raymond, W. D., Brown, E. L., & **Healy, A. F.** (2016). Cumulative context effects and variant lexical representations: Word use and English final t/d deletion. *Language Variation and Change, 28*(02), 175–202. DOI: https://doi.org/10.1017/S0954394516000041

Roberts, J. (2006). As old becomes new: Glottalization in Vermont. *American Speech, 81*(3), 227–249. DOI: https://doi.org/10.1215/00031283-2006-016

Selkirk, E. O. (1972). *The phrasal phonology of English and French.* PhD thesis, MIT.

Seyfarth, S. (2014). Word informativity influences acoustic duration: Effects of contextual predictability on lexical representation. *Cognition, 133*(1), 140–155. DOI: https://doi.org/10.1016/j.cognition.2014.06.013

Seyfarth, S., & **Garellek, M.** (2015). Coda glottalization in American English. In *Proceedings of the International Congress of Phonetic Sciences 2015*.

Strycharczuk, P., & **Scobbie, J. M.** (2016). Gradual or abrupt? The phonetic path to morphologisation. *Journal of Phonetics, 59*, 76–91. DOI: https://doi.org/10.1016/j.wocn.2016.09.003

Sumner, M., & **Samuel, A. G.** (2005). Perception and representation of regular variation: The case of final /t/. *Journal of Memory and Language, 52*(3), 322–338. DOI: https://doi.org/10.1016/j.jml.2004.11.004

Tanner, J., Sonderegger, M., & **Wagner, M.** (2017). Production planning and coronal stop deletion in spontaneous speech. *Laboratory Phonology: Journal of the Association for Laboratory Phonology, 8*(1), 1–39. DOI: https://doi.org/10.5334/labphon.96

Trager, G. L. (1942). The phoneme 'T': A study in theory and method. *American Speech, 17*(3), 144–148. DOI: https://doi.org/10.2307/486786

Trudgill, P. (1974). *The social differentiation of English in Norwich.* Cambridge: Cambridge University Press.

Vitevitch, M. S., & **Luce, P. A.** (1998). When words compete: Levels of processing in perception of spoken words. *Psychological Science, 9*, 325–329. DOI: https://doi.org/10.1111/1467-9280.00064

Wickham, H., Francois, R., Henry, L., & **Müller, K.** (2017). *dplyr: A grammar of data manipulation.* Retrieved from https://CRAN.R-project.org/package=dplyr

Yuan, J., & **Liberman, M.** (2011). Automatic detection of g-dropping in American English using forced alignment. In *2011 IEEE Workshop on Automatic Speech Recognition & Understanding.* IEEE. DOI: https://doi.org/10.1109/ASRU.2011.6163980

Zhao, S. Y. (2010). Stop-like modification of the dental fricative /ð/: An acoustic analysis. *The Journal of the Acoustical Society of America, 128*(4), 2009–2020. DOI: https://doi.org/10.1121/1.3478856

Zue, V. W., & **Laferriere, M.** (1979). Acoustic study of medial /t,d/ in American English. *Journal of the Acoustical Society of America, 66*(4), 1039–1050. DOI: https://doi.org/10.1121/1.383323