# The role of fast speech in sound change

**Uriel Cohen Priva (uriel_cohen_priva@brown.edu)**
Brown University
Providence, RI 02912 USA

**Emily Gleason (emily_gleason@alumni.brown.edu)**
Brown University
Providence, RI 02912 USA

## Abstract

Recent research has seen a surge in interest in the role of the individual in sound change processes. Do fast speakers have a unique role in sound change processes? Fast speech leads to greater rates of lenition (reduction). But should it mean that fast talkers would be more likely to lenite even when speaking slowly? In two corpus studies we show that even when fast talkers speak more slowly they are (a) more likely to omit segments and (b) more likely to perform variable reduction of consonants. This draws attention to habitual speech rate as a likely factor in the actuation of lenition processes.

keywords: lenition, speech rate, individual differences

## Introduction

A growing field of research focuses on social factors that may contribute to language change (Labov, 2001; Milroy & Milroy, 1985), including the influence of the individual. Studies of individual difference tend to revolve around different cognitive aspects of speakers, including position on the autism spectrum (Yu, 2013), attention to linguistic variation within and outside of personal linguistic subgroups (Garrett & Johnson, 2013), and differences in perception (Ladd et al., 2013). Other studies correlate the physiological makeup of speakers as a bias toward particular sound change processes (Moisik & Dediu, 2015), such that some properties of speech could be acquired by individuals and become typical of their language use. Baker, Archangeli, and Mielke (2011) proposes that it is differing amounts of variability between speakers which leads to change when an accepted variation by one group is misinterpreted as a new target by another group.

Individual speakers tend to be consistent in their behavior. In a study that focused on convergence in speech rate (Cohen Priva, Edelist, & Gleason, 2017), the correlation between the speech rate of a speaker in a given conversation and their speech rate in other conversations was very high ($\beta$=.79, predictors and predicted values were standardized). In contrast, speakers' speech rate was only mildly correlated (convergent) with the speech rate of their interlocutors in other conversations ($\beta$=.05). Though speakers do converge with their interlocutors in speech rate, they are more consistent than convergent or random. Similarly, high self-persistency was found in Cohen Priva and Sanker (2018) for other phonetic properties such as median pitch (self $\beta$=.97 vs. other $\beta$=0.018), pitch variability (self $\beta$=.68 vs. other $\beta$=0.092), and non-phonetic properties such as *uh* to *um* ratio (self $\beta$=.79 vs. other $\beta$=.031).[1] Speakers are therefore expected to "do what they always do", even when the local context demands otherwise (e.g. to converge). Exemplar-based representation (e.g. Pierrehumbert, 2001) could explain consistency if speakers sample from their own performance, which has many instances of their behavior, biased in a particular direction. Alternatively, consistency can follow from storage vs. computation considerations (O'Donnell, 2015) if speakers do not always compute the appropriate output, but instead sometimes reuse precomputed stored instances.

Regardless of the actual underlying mechanism, consistency in speech rate has surprising predictions for fine-grained phonetic behavior: it suggests that fast speakers would repeat "fast speech patterns" even when speaking more slowly. Fast speech has been argued to predict higher rates of lenition (reduction[2]) at the local context, e.g. in Kirchner (1998); Lavoie (2001); Gurevich (2004); Cohen Priva (2015). This implies that fast speakers will lenite more than slower speakers, everything else being equal. If fast speakers are repeating their fast speech patterns outside of the typical context, they would also lenite when speaking more slowly, which would implicate them as possible leaders in lenition-type sound changes. Theories regarding individual differences in sound change such as Yu (2013) assume that actuation can follow when listeners fail to interpret a particular exemplar as following only from the context in which it appears. Listeners observing lenition in the context of fast speech can attribute such effects to the speed the speech was produced. If lenited forms are produced out of context, when fast speakers happen to speak more slowly, they may not be attributed to fast speech, which could then follow the pattern discussed in Yu (2013). In such cases fast speakers may function as initiators and promoters of new and existing lenition-type sound change processes.[3]

We used the Buckeye corpus (Pitt et al., 2007) in several related studies to investigate the impact of fast speakers and

[1] For all reported $\beta$ values in Cohen Priva and Sanker (2018), predicted values and both predictors were standardized.

[2] Lenition is a poorly defined term surrounded by a signifcant amount of debate (see e.g. Honeybone, 2008), but we use it here as a short-hand for generally accepted reduction processes, i.e. degemination, voicing, spirantization, debuccalization, approximantization, flapping, and deletion.

[3] Here we do not mean to refer to stable and established lenitions (such as e.g. tapping in intervocalic contexts in American English), but rather new or ongoing sound changes that involve reduction.

fast speech on non-phonologized obstruent lenition. First, we investigated deletion, which is evident in the absence of underlying segments in the surface form, and has been used in the past in the Buckeye corpus (Cohen Priva, 2008, 2015, 2017a; Raymond, Dautricourt, & Hume, 2006). We then investigated lenition more broadly using a power difference estimation (Warner & Tucker, 2011). Both studies indicate that fast speakers do have a greater tendency to reduce segments, even beyond the effects of the speech rate within the individual phrase. Our findings thus highlight the complex role of individual speech rate in reduction processes.

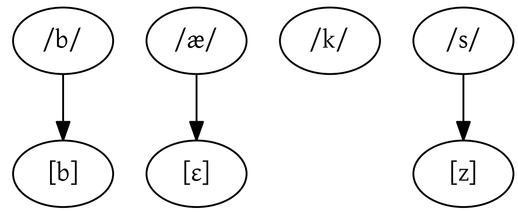


Figure 1: An example of the alignment of an underlying form with its surface representation

## Method and materials

### Corpus

For all studies below we used the Buckeye Corpus of Conversational Speech (Pitt et al., 2007). The Buckeye Corpus is a detailed corpus of American English composed of data from 40 speakers conversing with one of two interviewers at Ohio State University. Only the speaker side of each conversation was recorded. A small amount of demographic information about the speakers was recorded, including gender and age. Age was recorded as a binary value (whether a speaker was older or younger than 40 years of age).

Following established procedure (Cohen Priva, 2015), phonemes were linked to their surface realization using a weighted edit-distance program, but unlike Cohen Priva (2015), the weights were learned from the data algorithmically using a variant of the EM algorithm. The algorithm minimized the perpelxity of aligning the entire Buckeye corpus by assigining probabilities to every deletion, insertion, and substitution operation. Each individual underlying form was aligned to its corresponding surface form by, in essence, choosing the most probable alignment.[4] For example, in a case where the word /bæks/ was produced as [bɜz], the algorithm would align /b/ with [b], /æ/ with [ɜ], /s/ with [z], and regard /k/ as deleted, as shown in Figure 1. Underlying representations were taken from the CMU dictionary (Weide, 2008), as the Buckeye provided forms do not include stress information, which has been previously shown to be relevant to duration and deletion rates (e.g. Lavoie, 2001). Words that occurred less than 4 times in the data were omitted to allow convergence.

Segments that were not aligned with an underlying form were considered to have been deleted. The intensity in each sound file was extracted from the corpus using Praat (Boersma & Weenink, 2008), and aligned to individual segments based on their beginning and ending durations in the corpus. This was done to compare how loud segments were relative to the preceding vowel.

### Speaker speech rate and phrase-level speech rate

We followed the procedures used in Cohen Priva (2017b) to measure speech rate. The goal of this procedure is to abstract away from contextual factors and the phonological makeup of words, and measure whether words are pronounced faster relative to their expected pronunciation. The *expected duration* of each word was defined as the prediction of a linear regression that used the mean duration of the word in every context and its contextual predictability (Bell, Brenier, Gregory, Girand, & Jurafsky, 2009; Jurafsky, Bell, Gregory, & Raymond, 2001). *Pointwise speech rate* was defined as the actual duration of a word instance, divided by its expected duration. Therefore, if a word's duration was predicted to be 250ms but was pronounced in 300ms, its pointwise speech rate would be 1.2 (slow), while if that word were pronounced in 200ms, its pointwise speech rate would be 0.8 (fast). Note that higher pointwise speech rates indicate *slower* articulation. Speakers' speech rates were then calculated as the geometric mean of the pointwise speech rates of all content words they used in a conversation (function words defined as the list of words returned by the function `stopwords` in R's tm package, Feinerer & Hornik, 2017). Function words were not included as they have been argued to be retrieved using a different mechanism than content words (Bell et al., 2009).

All studies contrasted the effects of average speaker speech rate with phrase-level speech rate. Phrase-level speech rate was defined as average pointwise speech rate of all words in the phrase, except the pointwise speech rate of the word for which lenition was predicted. If a segment had been deleted from the word in question it would be trivially shorter, a confound between lenition and short duration. By considering both individual and phrase level speech rate we were able to compare how relevant an individual's overall speech pattern was for increased lenition compared to the previously-observed effect of fast speech within a phrase or word itself. For each model the interaction between average speaker speech rate and phrase rate was considered as a variable but did not have any significant effect on the results, and therefore it is not included in the models reported here.

Speech rate has been shown to depend on social factors, such as age and gender (e.g. Cohen Priva, 2017b; Jacewicz, Fox, & Wei, 2010; Quen, 2008; Yuan, Liberman, & Cieri, 2006). In data taken from the Switchboard and Buckeye cor-

[4] A side-effect of this process is that mergers were analyzed as a deletion of the segment less likely to emerge as the aligned surface form.

pora, men were found to generally speak faster than women (Cohen Priva, 2017b), despite the complete absence of evidence that would indicate that men are better than women in language use. Age and gender are therefore included as controls in the following studies.

## Study 1

For Study 1, we studied the effect of speech rate on obstruent deletion, a non-phonologized lenition process in American English. Deletion of a segment is easily detected through the alignment of segments to their output (or lack thereof, in the case of a deleted segment) as described above in the description of the corpus and its alignment. In Study 1a, we looked specifically at intervocalic environment, which is typical for lenition. In Study 1b, we extended the result to post-vocalic (pre-consonantal) environment.

### Study 1a: Intervocalic deletion

**Materials and methods** Log odds of deletion were predicted by a mixed effects logistic regression using ˜20,500 (intervocalic) segments. Log phrasal speech rate and log speaker average speech rate were both variables of interest. Age (binary), and gender were used as fixed effects. Segment, word, the phonological environment of the segment (no stress, secondary stress, or primary stress on either vowel), and speaker were included as random intercepts. All models presented here and below were fitted in R (R Core Team, 2017) using the lmerTest package (Kuznetsova, Bruun Brockhoff, & Haubo Bojesen Christensen, 2016), which encapsulates the lme4 package (Bates, Mächler, Bolker, & Walker, 2015).

**Results** Age and gender had no effect and were removed from the model (model comparison $p>.21$). Faster phrasal speech rate was indeed correlated with higher propensity to delete ($\beta$=-0.27, SE=0.04, z=-6.725, $p<10^{-10}$), consistent with current work on lenition.[5] Fast individual-level speech rate also correlated with higher propensity to delete ($\beta$=-0.32, SE=0.14, z=-2.215, $p<0.03$), suggesting an influence of the individual on lenition rates. The results are not due to collinearity between the two factors, as models that included an interaction term between phrasal and individual speech rate were not significantly different (or better) than the model provided here, and did not affect the significance of the individual-level speech rate.

**Discussion** High degrees of segment-deletion were known to correlate with fast speech (e.g. Cohen Priva, 2015), and many authors suggested that articulatory speed is a cause for lenition, perhaps by increasing the effort required by speakers (Kirchner, 1998), or by increasing the chances for hypoarticulation (Bauer, 2008). However, the results of this study do not have to follow from previous findings: Speech rate is variable, and speakers could have omitted segments only when speaking quickly, and preserved them when speaking more slowly. The evidence that fast talkers do omit segments even when speaking slowly could be explained if fast speakers have a speaking style that facilitates fast speech, and they continue to use it even when speaking more slowly. Causal direction here is only implied: an alternative is that perhaps being frequent omitters is the reason why they speak fast. Study 1b replicates this study in an additional environment.

[5] The coefficient is negative here because a larger value for speech rate indicates *slower* pronunciation.

### Study 1b: Post-vocalic deletion

**Materials and methods** Log odds of deletion were predicted by a mixed effects logistic regression using ˜15,600 (post-vocalic, pre-consonantal) segments. Since the following environment in this case was always a consonant, stress in the following environment was not included as a random intercept, but the manner of articulation of the following consonant was. All other variables and controls were the same as Study 1a.

**Results** Age and gender were not significant in this model as well, and were removed following a model comparison (model comparison $p>.5$). Phrase rate continued to be a significantly correlated with deletion in post-vocalic environment ($\beta$=-0.3, SE=0.047, z=-6.285, $p<10^{-9}$). Speaker rate, however, did not have a significant effect ($\beta$=-0.087, SE=0.1, z=-0.843, p=0.399). Models that included an interaction term between phrasal and individual speech rate were not significantly different (or better) than the model provided here.

**Discussion** While speakers were more likely to omit segments when speaking quickly, fast speakers were not more likely to omit segments in post-vocalic positions when speaking more slowly. It is possible that some reduction occurs, but does not culminate in outright deletion in this environment. Study 2a-b therefore focuses on more variable lenition types.

## Study 2

Although deletion is an easily detected lenition process, it is also possible that fast speakers play a role in the reduction of segments to a less severe degree. To capture this, we estimated reduction using the difference in power between the maximum of the previous segment and the minimum of the segment under investigation. This has two advantages. The first is that, as noted, this measurement is able to capture partial reductions rather than only outright deletions. The second is that the measurement is continuous rather than categorical. Lenition following modulation of intensity was discussed in Kingston (2008) and Katz (2016), and the procedure we use was inspired by Warner and Tucker (2011). The linear regressions in Study 2 predict the power differences for obstruents in American English. A greater power difference indicates a less lenited, more reduced, or more consonant-like segment. Smaller power differences indicate that the obstruent in question is closer in power to the vowel preceding it. This approach is more sensitive to variable lenition than a binary measurement (i.e. lenited or not), as it can capture smaller phonetic differences in articulation. Study 2a focuses on intervocalic environment, while Study 2b focuses on post-vocalic, pre-consonantal environment.

### Study 2a: Intervocalic lenition

**Materials and methods** Power differences were predicted by a mixed effects linear regression using ˜19,300 intervocalic segments (fewer than in Study 1a as deleted segments are not usable). All other variables and controls were the same as Study 1a.

**Results** Gender showed a significant effect, such that male speakers showed greater reduction in speech ($\beta$=-1.384, SE=0.53, t=-2.614, $p<0.05$). Younger speakers did not show a signiicantly different pattern of reduction than older speakers ($\beta$=-0.86, SE=0.52, t=-1.655, p=0.107). As expected, faster phrasal speech rate was correlated with greater amounts of reduction ($\beta$=0.37, SE=0.047, t=7.875, $p<10^{-14}$). Faster individual speech rate was also correlated with greater amounts of reduction ($\beta$=0.91, SE=0.25, t=3.672, $p<0.001$), following the pattern of Study 1a. Unlike Study 1a, the *effect size* was greater for individual speech rate than for phrasal speech rate.

**Discussion** These results strengthen the findings in Study 1a: Fast speakers are more likely to variably lenite and not only to omit segments even when speaking slowly. This lends support to the possibility that speakers are not fully flexible: speaking fast could affect the way speakers perform regardless of actual speech rate.

### Study 2b: Post-vocalic lenition

**Materials and methods** Power differences were predicted by a mixed effects linear regression using ˜14600 post-vocalic pre-consonantal segments. All other variables and controls were the same as Study 1b.

**Results** Male speakers were marginally more likely to reduce segments ($\beta$=-1.039, SE=0.59, t=-1.769, $p<0.1$). Younger speakers were significantly correlated with greater reduction of segments ($\beta$=-1.505, SE=0.58, t=-2.605, $p<0.05$). Faster phrasal speech rate was correlated with greater amounts of reduction ($\beta$=0.6, SE=0.056, t=10.698, $p<10^{-15}$), as in Study 1b. Unlike Study 1b (and like Study 1a and Study 2a), faster individual speech rate was also correlated with greater amounts of reduction ($\beta$=1.12, SE=0.27, t=4.171, $p<0.001$), and, like Study 2a, the effect size was greater than that of phrasal speech rate.

**Discussion** The results of Study 2b make it more likely that the absence of an effect for individual speech rate in Study 1b was not because fast speakers do not lenite more in post-vocalic environment, but such lenition processes do not necessarily culminate in outright deletion.

## General discussion

One general result of this paper is the corroboration of previous lab-based findings that highlight the difference between casual and careful speech (Warner & Tucker, 2011). This paper provides an important extension of these results, and correlates variable lenition (as measured in power difference) with speech rate in a large detailed corpus and not only in lab settings.

Our findings demonstrate that the typical speech rate of an individual does affect their behavior outside of known local effects. For all models, following the assumptions of most work in lenition, faster speech rate was found to be significantly correlated with increased rates of reduction. Moreover, in three of the four models (excepting Study 1b), the effect went beyond phrasal speech rate, the immediate rate around the word which contained a deletion, and extended to overall individual speech rate. In Study 1 this was shown for the deletion rate of intervocalic oral stops, a process not phonologically licensed in English, while in Study 2 it was replicated with power differences, a more nuanced measurement of reduction. We therefore conclude that fast speakers show *consistency* in their lenition patterns, through greater reduction in their speech, over and above the effects of local speech rate.

Importantly, the results draw attention to the inflexible properties of human speech. Cohen Priva (2017b) shows that fast speakers are more likely to use frequent words and less likely to use infrequent syntactic structures such as passive voice (without implying causal directionality). Though fast speakers could (and probably do) slow down when producing infrequent words and structures as suggested by several studies (e.g. Bell et al., 2009), they seem to use them less often. Speakers can and do change their speech rate while communicating with other speakers, but our results show they are also likely to adopt typical behaviors compatible with their overall performance. Perhaps fast talkers get accustomed to leniting articulatory patterns and apply them elsewhere, or, alternatively, the adoption of leniting articulatory patterns makes some speakers fast speakers. At least two mechanisms could support such patterns, as suggested above. First, speakers may sample from their own production and operate on sampled values before computing the contextual modification: if their own exemplar clouds have a high proportion of exemplars created in fast speech conditions, they are more likely to use them as a basis for subsequent productions. Second, in storage vs. computation frameworks such as O'Donnell (2015), speakers sometimes compute the most appropriate behavior, but sometimes use precomputed stored values: for fast speakers these are more likely to be lenited forms.

Fast talkers omit more and lenite more than slower speakers regardless of how fast they speak in an individual instance. For more nuanced reduction (Study 2), this effect coincided with the additional influence of gender and age, demographic variables typically found to influence sound change patterns. We conclude that fast speakers show a different pattern of reduction than slower speakers, and we hypothesize based on this evidence that fast talkers may play a critical role in the actuation of lenition-type sound change. Fast speakers may be the first to consistently provide lenited forms in non-fast speech (outside the "natural environment" for lenited forms), or perhaps their increased use of lenited tokens drives the phonologization process. This could have its own interesting implications for lenition-driven sound change and sound change by other means, such as mergers. In our results, male speakers trended to greater reduction, which goes against typical evidence that women and younger people are often the innovators of change (Labov, 2001). The result is consistent, however, with evidence that male speakers of American English tend to have faster speech rates than women, everything else being equal (Cohen Priva, 2017b; Cohen Priva et al., 2017).

To sum, this paper offers two conclusions that should affect future research in sound change and in the understanding of speech. First, it draws attention to fast speakers as likely initiators of lenition-type processes, as they are more likely to produce lenited forms overall but particularly in slow speech. Furthermore, our results show that even lenition-type behavior can be persistent, and follow from other properties such as speech rate.

## References


Baker, A., Archangeli, D., & Mielke, J. (2011). Variability in American English s-retraction suggests a solution to the actuation problem. *Language Variation and Change*, *23*(03), 347–374. doi: 10.1017/S0954394511000135

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: 10.18637/jss.v067.i01

Bauer, L. (2008). Lenition revisited. *Journal of Linguistics*, *44*(3), 605–624. doi: 10.1017/S0022226708005331

Bell, A., Brenier, J., Gregory, M., Girand, C., & Jurafsky, D. (2009). Predictability effects on durations of content and function words in conversational English. *Journal of Memory and Language*, *60*(1), 92–111.

Boersma, P., & Weenink, D. (2008). *Praat: Doing phonetics by computer (Version 5.0.26) [Computer program].* (Retrieved June 16, 2008, from http://www.praat.org/)

Cohen Priva, U. (2008). Using information content to predict phone deletion. In N. Abner & J. Bishop (Eds.), *Proceedings of the 27th west coast conference on formal linguistics* (pp. 90–98). Somerville, MA: Cascadilla Proceedings Project.

Cohen Priva, U. (2015). Informativity affects consonant duration and deletion rates. *Laboratory Phonology*, *6*(2), 243–278. doi: 10.1515/lp-2015-0008

Cohen Priva, U. (2017a). Informativity and the actuation of lenition. *Language*, *93*(3), 569–597. doi: 10.1353/lan.2017.0037

Cohen Priva, U. (2017b). Not so fast: Fast speech correlates with lower lexical and structural information. *Cognition*, *160*, 27–34. doi: 10.1016/j.cognition.2016.12.002

Cohen Priva, U., Edelist, L., & Gleason, E. (2017). Converging to the baseline: Corpus evidence for convergence in speech rate to interlocutor's baseline. *The Journal of the Acoustical Society of America*, *141*(5), 2989–2996. doi: 10.1121/1.4982199

Cohen Priva, U., & Sanker, C. (2018). Distinct behaviors in convergence across measures. In *Proceedings of the 40th annual conference of the cognitive science society.* Austin, TX.

Feinerer, I., & Hornik, K. (2017). tm: Text mining package [Computer software manual]. (R package version 0.7-3)

Garrett, A., & Johnson, K. (2013). Phonetic bias in sound change. In A. C. L. Yu (Ed.), *Origins of sound change: Approaches to phonologization* (pp. 51–97). Oxford, UK: Oxford University Press.

Gurevich, N. (2004). *Lenition and contrast : the functional consequences of certain phonetically conditioned sound changes*. New York: Routledge.

Honeybone, P. (2008). Lenition, weakening and consonantal strength: tracing concepts through the history of phonology. In *Lenition and Fortition.* Berlin, Boston: De Gruyter. doi: 10.1515/9783110211443.1.9

Jacewicz, E., Fox, R. A., & Wei, L. (2010). Between-speaker and within-speaker variation in speech tempo of American English. *The Journal of the Acoustical Society of America*, *128*(2), 839–850. doi: 10.1121/1.3459842

Jurafsky, D., Bell, A., Gregory, M. L., & Raymond, W. D. (2001). Probabilistic relations between words: Evidence from reduction in lexical production. In J. L. Bybee & P. Hopper (Eds.), *Frequency and the emergence of linguistic structure* (pp. 229–254). Amsterdam: Benjamins.

Katz, J. (2016). Lenition, perception and neutralisation. *Phonology*, *33*(1), 43–85. doi: 10.1017/S0952675716000038

Kingston, J. (2008). Lenition. In *Selected proceedings of the 3rd conference on laboratory approaches to Spanish phonology* (pp. 1–31).

Kirchner, R. M. (1998). *An effort-based approach to consonant lenition*. Unpublished doctoral dissertation, University of California Los Angeles.

Kuznetsova, A., Bruun Brockhoff, P., & Haubo Bojesen Christensen, R. (2016). lmertest: Tests in linear mixed effects models [Computer software manual]. (R package version 2.0-33)

Labov, W. (2001). *Principles of linguistic change, volume 2, social factors*. Oxford: Blackwell.

Ladd, D. R., Turnbull, R., Browne, C., Caldwell-Harris, C., Ganushchak, L., Swoboda, K., ... Dediu, D. (2013). Patterns of individual differences in the perception of missing-fundamental tones. *Journal of Experimental Psychology: Human Perception and Performance*, *39*(5), 1386.

Lavoie, L. M. (2001). *Consonant strength: Phonological patterns and phonetic manifestations*. Psychology Press.

Milroy, J., & Milroy, L. (1985). Linguistic change, social network and speaker innovation. *Journal of Linguistics*, *21*(2), 339–384.

Moisik, S. R., & Dediu, D. (2015). Anatomical biasing and clicks: Preliminary biomechanical modelling. In *18th international congress of phonetic sciences satellite event: The evolution of phonetic capabilities: Causes constraints, consequences* (pp. 8–13).

O'Donnell, T. J. (2015). *Productivity and reuse in language: A theory of linguistic computation and storage*. MIT Press.

Pierrehumbert, J. (2001). Exemplar dynamics: Word frequency, lenition and contrast. In J. Bybee & P. Hopper (Eds.), *Frequency and the emergence of linguistic structure* (pp. 137–157). John Benjamins Publishing Company.

Pitt, M., Dilley, L., Johnson, K., Kiesling, S., Raymond, W., Hume, E., & Fosler-Lussier, E. (2007). *Buckeye corpus of conversational speech (2nd release).* Department of Psychology, Ohio State University.

Quen, H. (2008). Multilevel modeling of between-speaker and within-speaker variation in spontaneous speech tempo. *The Journal of the Acoustical Society of America*, *123*(2), 1104–1113. doi: 10.1121/1.2821762

R Core Team. (2017). R: A language and environment for statistical computing [Computer software manual]. Vienna, Austria.

Raymond, W. D., Dautricourt, R., & Hume, E. (2006). Word-medial /t,d/ deletion in spontaneous speech: Modeling the effects of extra-linguistic, lexical, and phonological factors. *Language Variation and Change*, *18*.

Warner, N., & Tucker, B. V. (2011). Phonetic variability of stops and flaps in spontaneous and careful speech. *The Journal of the Acoustical Society of America*, *130*(3), 1606–1617. doi: 10.1121/1.3621306

Weide, R. (2008). *The CMU pronunciation dictionary, release 0.7a.* (Carnegie Mellon University)

Yu, A. C. L. (2013). Individual differences in socio-cognitive processing and the actuation of sound change. In A. C. L. Yu (Ed.), *Origins of sound change: Approaches to phonologization* (pp. 201–227). Oxford, UK: Oxford University Press.

Yuan, J., Liberman, M., & Cieri, C. (2006). Towards an integrated understanding of speaking rate in conversation. In *Proceedings of interspeech* (p. 541-544). Pittsburgh, PA.