CrossMark

Language Variation and Change, **28** (2016), 175–202.
© Cambridge University Press, 2016 0954-3945/16
doi:10.1017/S0954394516000041

Cumulative context effects and variant lexical representations: Word use and English final t/d deletion

WILLIAM D. RAYMOND, ESTHER L. BROWN AND ALICE F. HEALY
University of Colorado, Boulder

ABSTRACT

Word production variability is widespread in speech, and rates of variant production correlate with many factors. Recent research suggests mental representation of both canonical word forms and distinct reduced variants, and that production and processing are sensitive to variant frequency. What factors lead to frequency-weighted variant representations? An experiment manipulated following context and word repetition for final t/d words in read, narrative English speech. Modeling the experimentally generated data statistically showed higher final-segment deletion in tokens followed by consonant-initial words, but no evidence of increased deletion with repetition, regardless of context. Deletion rates were also higher the greater a word's cumulative exposure to consonant contexts (measured from distributional statistics), but there was no effect of word frequency. Token effects are interpreted in terms of articulation processes. The type-level context effect is interpreted within exemplar and usage-based models of language to suggest that experiences with word variants in contexts register as frequency-weighted representations.

Psycholinguistic studies of speech production have recently begun to investigate the representation and cognitive processing of variant word forms and have provided evidence that deleted variants, that is, words produced with canonical segments unpronounced, are represented in the mental lexicon, for at least some classes of words (e.g., Bürki, Ernestus, & Frauenfelder, 2010; Bürki & Gaskell, 2012; Hanique, Ernestus, & Schuppler, 2013), in addition to their canonical counterparts. The results of these production studies parallel the conclusions of speech processing research, which also suggest the mental representation of multiple lexemes for words (e.g., Connine, 2004; Pinnow & Connine, 2014; Pitt, Dilley, & Tat, 2011).

Given the likely existence of deleted variants in memory and the many variables associated with segment deletions (see, e.g., Brown & Raymond, 2014; Bürki, Ernestus, Gendrot, Fougeron, & Frauenfelder, 2011; File-Muriel & Brown, 2011; Guy, 1980; Raymond, Dautricourt, & Hume, 2006; Torreira & Ernestus,

We would like to acknowledge the valuable contributions of Reinhold Kliegl to the initial analyses of the experimental data using R. An earlier version of this report was presented on March 15, 2014, at the Georgetown University Round Table: Usage-Based Approaches to Language, Language Learning, and Multilingualism.

https://doi.org/10.1017/S0954394516000041 Published online by Cambridge University Press

2011; Zimmerer, Scharinger, & Reetz, 2011, 2014), one important question is, which of the many factors associated with production variability could lead to variant representations in the lexicon? The current study addresses this question in the context of the deletion of word-final /t/ and /d/ (-t/d) segments in American English. As an example, the word *fast* may be pronounced as [fæst] or [fæs]. Based on the results of the study, we propose that variant lexical representations of a word arise as a result specifically of speakers' previous experiences with a word in discourse contexts that promote variation. Word-boundary variation phenomena such as -t/d deletion readily allow for the investigation of contextual phonetic variability on production variation, because the -t/d words can be followed in speech by other words with a diversity of forms, or by a pause in speech, and different classes of contexts are associated with different variants.

Crucial to the study of variable contexts, we define a lexically specific measure of the likelihood that a word has occurred in following phonological contexts known to encourage the deletion of the final segment, specifically, in the discourse context of a following consonant-initial word, as in the sequence *fast course*. The likelihood measure is operationalized as the proportion of corpus occurrences of a word in the context promoting deletion (i.e., followed by an orthographically consonant-initial word) out of all occurrences of that word in the corpus. Equivalently, this is a measure, based on previous word use history, of the predictability of the deleting context given a word.

This measure instantiates for the -t/d deletion phenomenon a variable of a class we call frequency in a reducing context (FRC). FRC measures were first introduced in studies of other synchronic and diachronic deletion phenomena in Spanish data (Brown, 2004; Brown & Raymond, 2012; Raymond & Brown, 2012). In the current study, an experiment provides data by manipulating the deleting context (following consonant-initial or following vowel-initial word) for English -t/d words. Prior experimental studies of t/d variability that consider following segment context have to date only been carried out in other languages (e.g., German verb-final t-deletion; Zimmerer et al., 2011, 2014) and for other English t/d variants (e.g., de Jong [1998], for flapping). The experimental data are used in a generalized linear mixed-effect model (GLMM) to test whether the likelihood of a word's prior occurrence in following consonant contexts, called FRC_{del}, predicts a word's likelihood of -t/d deletion *independent of its following segment production context* in the experiment. An effect on realization of the cumulative, lexically specific, discourse context-dependent variable allows, it will be argued, an explanation for the origin of deleted variant representations within usage-based (e.g., Bybee, 2002) or exemplar (e.g., Pierrehumbert, 2001) models of language.

Deletion of -t/d is examined in this study because it is a common and well-studied variation phenomenon in American English speech (Guy, 1980; Labov, 1972; Neu, 1980; inter alia). A long history of -t/d variation study has established that the likelihood of -t/d deletion is conditioned by a word's phonological and morphological structure, by its context of production, and by

https://doi.org/10.1017/S0954394516000041 Published online by Cambridge University Press

sociolinguistic factors. Importantly, more recent work has pointed to effects on deletion rates related to aspects of cumulative language experience, as measured by the frequencies, probabilities, or predictabilities of words, of their constituents, or of sequences of linguistic units (e.g., Bybee, 2002; Jurafsky, Bell, Gregory, & Raymond, 2001a, 2001b).

One way to characterize -t/d deletion factors is that some are invariant across productions of a word, such as lexical structure, morphological composition, and grammatical class, but others vary across tokens of a given word, depending on the word's production context. Invariant factors each have the same potential impact on word tokens with that factor. For example with other variables held constant, rates of deletion for words in which the segment preceding -t is also a consonant (e.g., *fast*, with /s/ preceding the -t), will be higher than for a word with a vowel preceding -t (e.g., *fat*). Other invariant factors that have been proposed include the degree of similarity of the preceding segment to the final segment (Guy, 1980; Guy & Boberg, 1997) and stress on the -t/d syllable (Fasold, 1972; Labov, 1972).

Variable conditioning factors may or may not apply to an utterance of a given word. They include discourse context, as well as speaker identity and parameters of speech production. A salient phonological variable factor for -t/d deletion in discourse is the word-external context following a -t/d word. A following consonant-initial word has been found to encourage deletion of -t/d (e.g., the deletion of -t in *went back*) over the rate for the same word in the context of a following vowel-initial word or a pause (e.g., the deletion of -t in *went away* or *went* at the end of an utterance), a result that appears to be robust in corpus studies across a number of dialects (Fasold, 1972; Neu, 1980; Patrick, 1991; Santa Ana, 1991; Wolfram, 1969). Other variable factors include speech rate and speech fluency, both of which can affect rates of deletion for -t/d (Guy, Hay, & Walker, 2008), as well as other reduction phenomena (Jurafsky, Bell, Fosler-Lussier, Girand, & Raymond [1998], for English function word reduction; Raymond et al. [2006], for English word-internal t/d deletion).

Some studies have shown that cumulative experience with individual words *independent* of context (i.e., word frequency) is correlated with rates of -t/d deletion and other reduction phenomena (e.g., Bybee, 2002; Coetzee, 2009; Guy et al., 2008; Jurafsky et al., 2001a, 2001b; Phillips, 2006). Research on the cumulative effect of words in contexts is more limited. In one study, Bybee (2002) showed that the rate of occurrence of some classes of -t/d words in a deleting context is correlated with -t/d deletion rate for the class, but she did not examine the effect of context rates for individual words. In a study of -t/d deletion using a corpus of New Zealand English, Guy et al. (2008) included a lexically specific but noncontinuous variable reflecting a word's most common following environment (deleting, nondeleting) and found that this variable predicts deletion.

Brown (2004) proposed a lexically specific type-level measure of the cumulative experience of a word in reducing contexts. Following this line of research, Raymond and Brown (2012) showed that the synchronic rate of initial

https://doi.org/10.1017/S0954394516000041 Published online by Cambridge University Press

/s/ deletion for a word in New Mexico Spanish is conditioned by the word's rate of occurrence in the context of a preceding nonhigh vowel, even after controlling for other invariant and variable factors. They also emphasize that word frequency has not been identified as contributing to reduction consistently within or across similar studies, which may at least partially be explained as a failure to also include FRC measures in corpus analyses.

The rate of a word's occurrence in the -t/d deleting context of a following consonant-initial word, which will be referred to as FRC_{del}, has not previously been tested as a conditioning factor in studies of -t/d deletion. Thus the primary goal of the current study is to test the effect of the continuous measure FRC_{del} on -t/d deletion. A GLMM of the experimental data will allow comparison of production context experience, FRC_{del}, and word frequency as predicting factors for -t/d deletion, while taking into account numerous other factors known to influence deletion rates, including differences among subjects and words used in the experiment.

Effects of contextual occurrence rates on variant production would have implications for psychological theories of language. Knowledge of word variant distributions across word use contexts is amenable to explanations in an exemplar or usage-based theory of language (Bybee, 2000, 2002; Pierrehumbert, 2001). We can additionally ask whether there is a short-range effect of a word's production context on its subsequent production; specifically, will the production of a -t/d word in a following consonant context (presumably the context encouraging deletion, as in other studies) increase the likelihood of deletion on the word's next production, regardless of the second production's context? In a corpus data of interview speech, Tamminga (2014) found no effect of a word's following phonological context on the likelihood of deletion on the subsequent production of the word. Note, however, that there was no consideration in her study of the cumulative effect of following context. Another goal of the present study, then, is to assess whether an effect of deleting context persists when words are repeated using the controlled setting of read speech that contains both repeated words and new words, with following phonological context manipulated for both.

EXPERIMENT

The experiment was designed to generate a corpus by having participants read -t/d words in two narrative paragraphs in which semantic context and numerous lexical variables that affect final segment deletion likelihood were controlled across lists of words. The two experimental paragraphs were constructed to manipulate the following phonological context in which target words occurred and whether target words were repeated across paragraphs. Words followed fluently by a consonant-initial word were expected to have higher deletion rates than words followed by vowel-initial words or with a pause following the -t/d word.

https://doi.org/10.1017/S0954394516000041 Published online by Cambridge University Press

Repeated words may have higher deletion rates than nonrepeated words when repeated words were initially produced in the deleting context.

METHOD

Participants

The participants consisted of 48 undergraduate students (33 female and 15 male) at the University of Colorado, Boulder, who participated in the experiment for credit in a course on introductory psychology during the fall 2012 semester. All participants but one were native speakers of English as their sole first language.[1] Half of the participants had grown up in Colorado; the remaining 24 had spent their childhoods in various other states in the East ($n = 6$), Midwest ($n = 9$), and West ($n = 9$). There were no speakers of Southern, African American, or any r-less varieties of English.

Materials and design

The target stimuli consisted of three sets of 10 words each (see Appendix A). The words all ended canonically in /t/ or /d/ (with /t/ represented orthographically by *t* and /d/ by *d* for all words except *crossed*). The properties of word length, stress pattern, word class, morphological constituency, phonological environment preceding the -t/d, and whether the word's final segment is /t/ or /d/ are all known from previous studies to affect the likelihood of -t/d deletion (for a detailed review see, e.g., Raymond et al., 2006). Words were loosely matched on these factors across the three lists, as shown in Appendix A. Distributional frequencies for the stimulus words, which have also been correlated with deletion rates, were determined using the Corpus of Contemporary American English (COCA) oral corpus (of about 96 million words; Davies, 2008–). Corpus counts were transformed to $\log_{10}$ per million words for use in analysis. Log transformed frequencies did not differ between lists ($F[2, 27] < 1$). FRC_{del}'s of stimulus words in following consonant contexts were also calculated using the COCA oral corpus (Davies, 2008–). For a given word *w*, FRC_{del} is defined as the number of occurrences of *w* in the context encouraging deletion (i.e., a following consonant-initial word) out of the total number of occurrences of *w*. Note that this definition makes FRC_{del} equivalent to the predictability of the deleting context given the word *w*. Average FRC_{del}'s did not differ among lists ($F[2, 27] < 1$).

To prepare the experimental materials, three narrative paragraphs were created, each based on a different narrative story. Each of the three paragraphs included the words from a different pair of the -t/d lists, and all were written so that each target word was followed by a consonant-initial word. (The three paragraphs will be referred to as P12C, P13C, and P23C, indicating their word list combinations and following context). The words from the two lists in each consonant paragraph were interleaved (but not strictly alternated) in the paragraphs' texts.

https://doi.org/10.1017/S0954394516000041 Published online by Cambridge University Press

A variety of consonants were used for the following context in each paragraph to avoid limiting any effect on deletion of a particular consonant.[2] There were no -t/d words in any of the paragraphs other than the target words. Three additional paragraphs (P12V, P13V, and P23V) were constructed from the three consonant context paragraphs by replacing consonant-initial following words by vowel-initial words, making any minor changes required to retain syntactic and semantic coherence in each narrative. The following vowel-initial words were chosen to begin with different vowels to avoid limiting any effect on deletion to a particular vowel quality. Corresponding C and V paragraphs (P12C and P12V; P13C and P13V; and P23C and P23V) were thus largely lexically, semantically, and syntactically equivalent, differing only in the identity of the words following the target -t/d words. To test the comparability of target word and following word pairs in C paragraphs and V paragraphs, target-plus-following-word bigram frequencies were compared across the three narrative pairs of C and V paragraphs. Word bigram frequencies were calculated by combining counts from the SUBTLEX-UK corpus (van Heuven, Mandera, Keuleers, & Brysbaert, 2014), the Buckeye corpus (Pitt, Dilley, Johnson, Kiesling, Raymond, Hume, & Fosler-Lusier, 2007), and the Switchboard corpus (Godfrey, Holliman, & McDaniel, 1992) (a total of about 205 million words). Word bigram frequencies per million words did not differ between paragraphs for any of the three list pairs ($F[2, 57] < 1$). An example stimulus paragraph is shown in Appendix A.

The C and V paragraphs were paired to counterbalance list combinations in each paragraph of a pair and the class of the segment following each -t/d word (C or V) in each paragraph of a pair. First, note that in order to test the effects of repetition, in each paragraph of a counterbalancing pair, one of the two lists included in the paragraph was required to occur in both the first and second paragraphs (the repeated list); the other list in each paragraph was required to be different from the repeated list and different from the nonrepeated list in the paired paragraph (the nonrepeated lists). For example, one paragraph pair consisted of P12C and P13C. The repeated list in this pair is list 1; list 2 and list 3 are nonrepeated lists. The result of the counterbalancing was a 3 (list combination in first paragraph: 1 and 2, 1 and 3, or 2 and 3) × 2 (repeated list: first or second list of the pair) × 2 (following context in first paragraph: C or V) × 2 (following context in second paragraph: C or V) design with 24 conditions. Across the conditions, each list is thus presented as repeated and nonrepeated, together with every other possible combination of alternate list. Also across the conditions, each paragraph, whether C or V, is paired with all possible combinations of other paragraphs, both C and V. This counterbalancing allowed control of stimulus words in terms of both repetition and C-V context.

Procedure

Paragraphs were presented to participants on the screen of a MacBook Pro in Arial 22-point font. On the screen no -t/d words occur at the end of a line, to encourage fluent production of target -t/d words and their following context words (referred to

https://doi.org/10.1017/S0954394516000041 Published online by Cambridge University Press

hereafter as target pairs). The first 24 participants were each placed in 1 of the 24 conditions; the second 24 participants were placed in the 24 conditions in the same condition order as the first 24. Participants initially read instructions indicating that they would be reading two paragraphs out loud and their read responses would be recorded using the MacBook Pro's internal microphone. The participants were also instructed that there would be a comprehension test following their reading of each paragraph. The tests were to encourage participants to focus on the meaning of the paragraphs and not individual words. After instructions were read, recording was initiated on the computer using the Audacity software program (Audacity Team, 2008), and the first paragraph was presented in its entirety. Participants read the paragraph at their own pace. At the end of each paragraph participants answered, at their own pace, four four-alternative multiple choice comprehension questions that were presented on the computer. On completing the comprehension test, the second paragraph was presented, followed by its comprehension test.

RESULTS

Responses to the comprehension tests averaged 85% correct (the modal number of correct answers was seven of eight on both comprehension tests), indicating that participants had generally understood the paragraphs they had read. Despite some reading errors on words throughout the paragraphs, almost all target pairs were read correctly; of the 1920 responses, only 57 target pairs (3.0%) were produced with errors. Errors included phonological and lexical speech errors on targets (e.g., "lifth up" for "lift up"; "brought enough" for "bought enough"), or cutoffs within target pairs without a fluent restart that included the target (e.g., "delayed f- because" for "delayed because"). The 57 target pairs produced with such errors were excluded from further analysis.

Coding

The 1863 nonerror target pair responses were examined by one of the authors (W.D.R.). Acoustic coding was conducted on each item using the phonetic analysis software Praat (Boersma & Weenink, 2005) and visual inspection of the sound wave and spectrogram displays. Each target token was coded as being in one of four categories: (a) *canonical:* -d words realized as [d] or -t words realized as [t]; (b) *flap*: targets realized with a final oral flap or nasal flap; (c) *glottal:* glottalization or glottal closure at the end of a target word in the absence of alveolar articulation; or (d) *deleted*: failure to meet the criteria for any of the previous variant realizations, so that there was no (or incomplete) acoustic evidence of a segment in the target word's production where the -t/d was expected. The specific criteria used for each of the coding categories are described in detail in the supplemental online appendixes (the online appendixes can be viewed at http://cambridge.journals.org/LVC), where spectrograms exemplifying each category in C and V contexts can also be found. Nonerror target tokens were also coded for fluency of their target pair. Pairs were coded as

https://doi.org/10.1017/S0954394516000041 Published online by Cambridge University Press

dysfluent if they were produced with a pause (and lack of a speech signal) separating the -t/d word from its following context word. Of the nonerror target pair responses, 150 target pairs (8.1%) were produced dysfluently.

Descriptive statistics

Table 1 shows the numbers of nonerror target realizations in each coding category for -t and -d words. Most of the target -t/d segments were either canonically realized (40.5%) or realized as a variant segment (glottalized: 15.2%; flapped: 9.9%), but 34.4% of the -t/d segments were deleted, indicating that deletion is a common phenomenon for these words in this read, English narrative speech (cf. Zue & Laferriere, 1979; Mitterer & Ernestus, 2006). Deletion rates in the experiment were higher than in corpus studies of spontaneous speech, because the words and contexts chosen for the experiment influenced deletion rates.

Noncanonical segment realizations were twice as likely to come from -t as from -d. The -t's were most likely to be realized as glottals and -d's as flaps. The higher rate of -d flapping is consistent with previous proposals that /d/ is more likely to flap than /t/ (Rhodes, 1992), but might also reflect a greater likelihood of transcription confusion between [d] and [ɾ] due to acoustic similarities and the gradient nature of flap articulation (de Jong, 1998). However, criteria for categorizing [d] and [ɾ] were developed and applied for transcription (see the online appendix). Finally -t's were more likely overall to delete than -d's (40.1% of -t's deleted and 26.0% of -d's deleted). This differing -t and -d deletion rate suggests a possible effect of final segment identity on deletion. However, -t and -d targets did not occur with the same set of following words, nor did they have the same set of preceding phones. In the raw data, the greater deletion rate for -t holds even if we compare deletions only in following phone contexts in which both -t and -d occurred (41.9% of -t's and 27.4% of the -d's deleted in these environments), but comparing deletion rates for -t and -d words with equivalent preceding segments results in similar deletion rates for -t (26.5%) and -d (28.7%).

TABLE 1. *Numbers of realizations for -t and -d target words in each coding category by production context (with percent of realization for each context)*

Canonical type	Production context	Realization				Total
		Canonical	Flap	Glottal	Deletion	
/d/	Consonant	206 (62)	19 (6)	1 (0)	104 (32)	330
	Vowel	99 (28)	146 (42)	15 (4)	90 (26)	350
	Pause	66 (98)	0 (0)	1 (2)	0 (0)	67
/t/	Consonant	131 (25)	3 (1)	107 (20)	280 (54)	521
	Vowel	212 (41)	17 (3)	118 (23)	165 (32)	512
	Pause	40 (48)	0 (0)	41 (49)	2 (2)	83
Total		754 (40)	185 (10)	283 (15)	641 (34)	1863

https://doi.org/10.1017/S0954394516000041 Published online by Cambridge University Press

Table 1 also shows that most of the dysfluent targets were produced as canonical [t] or [d] (71% of all dysfluent responses). Although dysfluent -d targets were almost always [d] (98%), dysfluent -t targets were much less likely to be [t] (48%). A high percentage of dysfluent [t] responses were glottalized (49%). Canonical [t] and [d] responses also had substantial percentages of following pauses (18% of all [d] responses and 10% of all [t] responses). Conversely, there were no pauses following flaps and only two target pairs had pauses following -t/d deletion. The high proportion of pauses after canonical realizations overall supports claims that a pause in speech discourages reduction in the word preceding the pause (e.g., Jurafsky et al., 1998; Wolfram, 1969). The lack of pauses after flaps reflects the fact that the flapping environment in English is prevocalic. The high percentage of -t glottalizations before pauses supports suggestions that pause is a glottalizing environment for -t (see Eddington & Channer, 2010). In fact, the percentage of pauses after glottalizations (15%) was similar to the percentages of pauses after canonical [t] (10%) and [d] (18%). The paucity of pauses after deletions suggests that deletion is, at least in part, conditioned by a following segmental context and motivates the exclusion of dysfluent tokens from the analyses of deletion.

Consistent with previous results, the rate of deletion is highest in _C environments (32% for /d/ and 54% for /t/). As expected, flapping preferentially occurred in _V environments, and flaps mostly derived from -d's (42%). Glottalization occurred often in _V environments, mostly from –t's (23%), although a substantial proportion of –t's in _C environments glottalized (20%). The high rate of glottals before vowels is unexpected from other studies of glottalizing, which have claimed that following consonants encourage glottalizing (see Eddington & Channer, 2010); however, the difference is perhaps explained by the restricted set of words and contexts used in the current study.

Dysfluent productions (i.e., before pauses) were fairly equally divided between intended segment contexts: 45% of pauses preceded C words. There was no difference by realization in the likelihood of a dysfluency to occur in an intended consonant or vowel context ([t]: consonant = 15, vowel = 25: [d]: consonant = 34, vowel = 32; glottalization: consonant = 17, vowel = 25, $\chi^2 = .896$; $p > .10$).

Table 2 shows numbers of each realization category by target word. As shown in Table 2, the words can be informally organized into three groups, determined by distribution of the words' realizations: (a) a *canonical* group, in which the majority of nondeletions (an average of 99% across words) were realized canonically as [t] or [d] (with only two flaps and no glottalizations); (b) a *flapping* group, in which the most frequent nondeletion realizations were flaps or canonical forms (an average of 97% across words) (with at least four flap realizations per word, but fewer than four glottalizations); and (c) a *glottalizing* group, in which the most frequent nondeletion realization was glottalization (an average of 77% across words).

Realization groups largely correlated with final segment type and preceding segment context. Words in the canonical group are -t words with a stop (/k/) or fricative (/f/, /s/) preceding the -t. Words in the flapping group are all -d words and have a vowel, /r/, or /l/ preceding the -d. Words in the glottalizing group are all -t

https://doi.org/10.1017/S0954394516000041 Published online by Cambridge University Press

TABLE 2. *Number of realizations in each coding category by target word, grouped by variant distributions*

Variant group	Word	Final segment realization			
		Canonical	Flap	Glottal	Deletion
Canonical	Fact	55	0	0	9
	React	51	0	0	11
	Act	49	0	0	15
	Craft	42	0	0	19
	Lift	30	2	0	25
	Gift	31	0	0	32
	Crossed	21	0	0	40
	Fast	20	0	0	44
	Best	14	0	0	47
Flapping	Yard	47	4	2	10
	Card	38	12	2	10
	Died	29	14	0	20
	Gold	35	10	0	18
	Seated	30	16	0	18
	Tired	28	11	1	22
	Wasted	23	16	1	20
	Wanted	24	13	2	22
	Board	23	20	2	17
	Said	32	15	2	13
	Delayed	36	17	2	8
	Avoid	26	17	3	16
Glottalizing	Apartment	0	0	28	33
	Engagement	1	0	28	34
	Went	2	1	16	44
	Bought	6	11	27	16
	Admit	7	2	50	5
	Regret	8	4	33	18
	Incident	11	0	30	21
	Fault	16	0	35	11
	Guilt	19	0	19	23

words and are preceded by /n/, /l/, or a vowel. The association of realization group with phonological structure is consistent with studies showing that -t/d variant frequencies are largely predicted by phonological context in corpus and experimental studies (e.g., Pitt et al., 2011; Raymond et al., 2006; Zue & Laferriere, 1979).

Unlike the patterns across groups of segmental realizations, words in all three groups had tokens produced with deleted final segments, and deletion averages in each group were substantial (canonical: 41%; flapping: 27%; glottalizing: 37%), with no significant differences among the groups ($F_2[1, 27] = 2.57$, MSE = .033, $p = .095$). Thus, although segment variant productions seem to be largely determined by lexical structure and following context, deletions are not as strongly associated with specific segment contexts. Nevertheless, segment context clearly played a role in deletion; the highest deletion rates are in words in which -t is preceded by /s/ or /n/ (54.0% to 77.0% deletion for the six highest deleting words: *best, fast, went, crossed, engagement, apartment*), as has been

https://doi.org/10.1017/S0954394516000041 Published online by Cambridge University Press

found in previous studies (e.g., Mitterer & Ernestus, 2006; Neu, 1980). Following context also appears to play a role in deletion, with tokens in consonant contexts deleting more (45%) than tokens in vowel contexts (30%). We thus expect to find effects of preceding and following contexts in the analysis of deletions. However, the commonality and variability of deletion within and across all word groups suggests that other word-specific factors may also affect deletion.

Finally, deletion rate also differed by participant gender and past tense morphology. Men deleted (50%) more than women (31%), suggesting a possible gender effect on deletion. There was a slightly higher deletion rate for past tense forms (38%) than non–past tense forms (37%), motivating inclusion of past tense morphology (past tense, other) as a factor in the analysis.

GLMM

The experiment generated a corpus of read, narrative speech containing tokens of a set of 30 -t/d words. The fluent, nonerror tokens from this corpus ($n = 1713$) were subjected to an analysis of variance (ANOVA),[3] and the complete results of the ANOVA are presented in the online appendix. Note that using ANOVA on data from an experimental design in which the dependent variable is an average over binary values (like proportion deleted in the current study) can be shown to be potentially misleading, and logistic regression cannot be used for repeated-measures designs (Dixon, 2008). As a solution, Dixon argued for a GLMM of data from repeated-measures designs (see also Kliegl, Wei, Dambacher, Yan, & Zhou, 2011). GLMMs avoid scaling effects from averaging and also allow examination of the random effects of item and subject variables on the binary response variable, as well as examination of lexical and subject factors not manipulated in the experiment. Thus, these experimental data were subjected to GLMM analyses in R (R Development Core Team, 2010) using the function *lmer()* with the binomial family specified (Bates, 2005; Bates, Maechler, & Bolker, 2011). (An analysis of all nonerror tokens, that is including the relatively small number of dysfluent tokens, revealed quite similar effects.)

Overall analysis

The independent variables considered in the GLMM included following context, list repetition, and paragraph, which allowed us to examine the effect of following context on -t/d production and the effect of following context on subsequent productions of -t/d words. Recall that following context was defined in terms of the initial segment of the word occurring after the target word. In order to test the second goal (to examine a cumulative effect of following production context), we included in the model our FRC_{del} measure (proportion of COCA occurrences of a word in _C contexts out of all contexts), again with the following production context used in the calculation of FRC_{del} defined in terms of the initial segment of the word that occurred in the corpus after the target word. Additional lexical variables that were also included in the regression analysis were word frequency (log frequency per million from COCA; Davies, 2008–), syllable stress (lexical

https://doi.org/10.1017/S0954394516000041 Published online by Cambridge University Press

stress on the -t/d syllable: stressed, unstressed), following syllable stress (lexical stress on the first syllable of the word following the target word: stressed, unstressed), segment identity (/t/, /d/), past tense morphology (past tense, other), and preceding context (i.e., the segment preceding the final -t/d within the target word). Two different binary categorizations of the preceding contexts of the target words were examined. The first distinguished consonants (consonant group) from vowels (vowel group). The second categorization distinguished the (oral) flapping and the nonflapping environments by grouping obstruents and /n/ (i.e., /f/, /k/, /n/, and /s/; the nonflapping group) against vowels, liquids (/l/), and glides (/r/) (the flapping group). Recall that only one token of a word with /n/ preceding -t was realized as a nasal flap, which motivated the inclusion of /n/ in the nonflapping group. Stress was not considered in defining the flapping environment, because stress on the following syllable may not be a conditioning factor for interword flapping (Oshika, Zue, Weeks, Neu, & Auerbach, 1975). The classification of preceding environments as flapping versus nonflapping proved better than consonant versus vowel in predicting deletion and was used in the final model. Contributions of two subject variables were also examined: gender (male, female) and paragraph speech time (PST; the total time taken by a participant to read an experimental stimulus paragraph). PST was used as a global measure of speech rate and fluency. PST ranged from 47.8 sec to 106.0 sec, with a median of 69.6 sec.

Initial examination of the dataset suggested that the continuous variables FRC_{del}, word frequency, and PST may interact with preceding context, following context, syllable stress, and gender. Variables were added to the model individually, with two-way interactions subsequently tested separately (see Dixon, 2008). Word frequency, PST, past tense morphology, and following syllable stress were not significant ($p > .05$) and did not participate in any significant interactions; these variables were not included in the final model as fixed effects. There were also no significant contributions from paragraph (first, second), list repetition, or segment identity when these variables were entered into the model, and so were also eliminated. Segment identity was not a significant predictor of deletion despite differences in the distribution of -t and -d realizations. However, -t and -d occurred in words with different sets of preceding segments. The confounding of preceding segment context and segment identity does not allow us to draw definitive conclusions about the differences between deletions of -t and -d.

In the model, subject and word were included as random effects (see Baayen, Davidson, & Bates, 2008). Random subject intercepts contributed significantly to the prediction of deletion (i.e., a likelihood ratio test of models with and without random subject intercepts indicated that the models were significantly different; $\chi^2 = 60.2$, $p < .001$); however, models with and without by-subject random slopes for FRC_{del} did not differ ($\chi^2 = .17$, $p = .92$), and thus the slopes were not included in the final model. Random word intercepts also contributed to model accuracy ($\chi^2 = 24.2$, $p < .001$), as did by-word random slopes for PST ($\chi^2 = 14.9$, $p < .001$).

https://doi.org/10.1017/S0954394516000041 Published online by Cambridge University Press

The significant final model results are summarized in Table 3. (Final model results including nonsignificant main effects and interactions can be seen in Appendix B.) Models were fitted using the restricted maximum likelihood method.[4] Neither the preceding nor following contexts showed a main effect; however, their interaction was significant. As seen in Figure 1, the interaction

TABLE 3. *Final model: Significant factors predicting -t/d deletion* (n = *1713*)

Parameters	*b*	SE	*p*
(Intercept)	−2.70	.86	.002
FRC_{del}	6.67	1.70	<.001
FRC_{del} × following context	−1.94	.77	.012
FRC_{del} × preceding context	−7.03	1.98	<.001
Following context × preceding context	.929	.25	<.001
Following context × syllable stress	1.39	.29	<.001
Gender × preceding context	.751	.26	<.003
Gender × syllable stress	1.03	.31	<.001

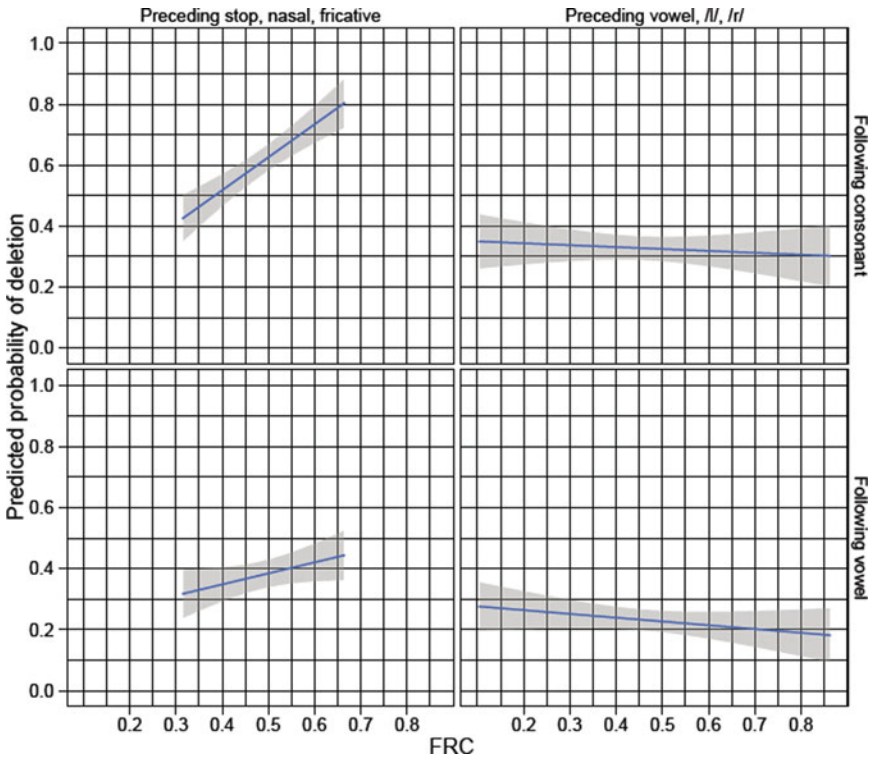


FIGURE 1. Model predictions of deletion for all preceding and following contexts as a function of word FRC_{del}.

https://doi.org/10.1017/S0954394516000041 Published online by Cambridge University Press

reflects more deletion in the nonflapping context than in the flapping context, with a greater effect when following context was a consonant than a vowel. The main effect of FRC_{del} reflects the fact that higher FRC_{del} words were more likely to have -t/d deleted than words that had lower FRC_{del}. However, FRC_{del} also separately interacted with preceding context and with following context. As shown in Figure 1, there was a strong positive correlation between FRC_{del} and deletion in the nonflapping context (shown in the left two panels of Figure 1), but, unexpectedly, no positive correlation in the alternative flapping context (shown in the right two panels of Figure 1). Additionally, in the nonflapping context, there was a stronger effect of increasing FRC_{del} in the context of a following consonant than of a following vowel.

The last four interactions from the model in Table 3 are summarized in Table 4. The interaction of following context and syllable stress reflects the fact that stress on the -t/d syllable made little difference when the following context was a consonant. Deletion rates were lowest when the following context was a vowel and there was lexical stress on the -t/d syllable, which encompasses the flapping context. Men and women responded differently to preceding context and syllable stress. Men showed more deletion across categories than women, but there was a greater difference between men and women when the preceding context was flapping than nonflapping, and also when the -t/d syllable was unstressed than when it was stressed. The less reduced speech of women is consistent with other findings in the literature that women use fuller, standard forms more than men (Labov, 1972), including deleting -t/d less in spontaneous speech (Byrd, 1994; Neu, 1980; Wolfram, 1969), but also for a range of other reduction phenomena cross-linguistically (e.g., lower rates of regressive voice assimilation in Dutch, Ernestus, Lahey, Verhees, & Baayen [2006], and -t/d deletion in German, Zimmerer et al. [2011]).

Analysis of flapping environment tokens

Surprisingly, the regression analysis of the data failed to find an increase in deletion with increasing FRC_{del} for tokens with preceding vowel, /l/, or /r/, as seen in the right two panels of Figure 1. When words in these preceding contexts are followed by a vowel-initial word (the lower right panel of Figure 1), they are

TABLE 4. *Percent deleted in the raw data for independent variables in model interactions (with cell* n*)*

		Gender		Following context	
		Female	Male	C	V
Preceding context	Nonflapping (/k, s, n, f/)	45 (510)	58 (244)	61 (375)	38 (379)
	Flapping (/V, l, r/)	21 (657)	44 (302)	33 (476)	23 (483)
Syllable stress	Stressed	31 (933)	46 (437)	46 (686)	26 (684)
	Unstressed	33 (234)	65 (109)	44 (165)	43 (178)

https://doi.org/10.1017/S0954394516000041 Published online by Cambridge University Press

susceptible to the flapping process, and flapping was a common realization for these tokens (30.1% of realizations in this environment were flaps); thus the history of their exposure to the flapping process was suspected to play a role in their behavior with respect to deletion as well. The prediction is additionally motivated because it is known that there is, for a variety of speech styles, a continuum of flap realizations exhibiting increasing reduction from stops to approximants to complete segment deletions (Warner & Tucker, 2011). Figure 1 shows a slightly negative correlation between deletion rate and FRC_{del} for the flapping context tokens. That is, there were somewhat more deletions for these tokens when flapping words had had *less* exposure to the deleting context, rather than more. Because flapping occurs in environments that are not deleting environments, a negative correlation with deletion history would suggest a positive correlation with flapping history. This observation suggests that flapping context words are, perhaps, more likely to undergo deletion as the consequence of increasing likelihood of flapping. A measure of flapping experience would be the rate of occurrence of these words in a flapping context (i.e., before vowel-initial words in context), called FRC_{flap}, which (excluding prepausal counts) is the complement of FRC_{del}. The effect of prior experience with flapping would be strongest when flapping is known to be most likely, that is, in stressed syllables followed by a vowel-initial word (see, e.g., de Jong, 1998), and might mean FRC_{flap}, rather than FRC_{del}, would predict deletion in a statistical model of these environments.

To test this hypothesis, a GLMM was fitted to the flapping context tokens from the deletion data in stressed syllables followed by a vowel context ($n = 391$). FRC_{flap} was calculated for the target words from the COCA oral corpus (Davies, 2008–) counts. As expected, a word's rate of occurrence in the deleting context (FRC_{del}) was fairly strongly negatively correlated with its rate of occurrence in the flapping context (FRC_{flap}) ($r^2 = -.522$). Following the procedure for the model of deletion in the complete dataset, FRC_{flap} was entered into the model, along with following context, gender, PST, and following syllable stress. Word intercepts and by-word slopes with respect to PST were included as random effects in the model.

The best model results of the analysis are seen in Table 5, which shows the effect of FRC_{flap} on deletion. As anticipated, the greater the likelihood that these words had appeared in the flapping context in COCA, the more likely they were to be realized as deleted. There was also an effect of gender, indicating that men

TABLE 5. *Final model predicting -t/d deletion for flapping contexts target tokens* (n = *391)*

Parameters	*b*	SE	*p*
Gender	1.36	.387	<.001
FRC_{flap}	2.83	1.14	.013

https://doi.org/10.1017/S0954394516000041 Published online by Cambridge University Press

deleted more than women in flapping environments, just as they had in all other preceding context and syllable stress subsets (see Table 4).

DISCUSSION

The analyses of the experimental data generated from read speech revealed both a token-level effect of discourse production context and a type-level effect of context history. How can these effects be explained in a model of language representation and production? The token effect of consonant context on deletion can be modeled as an on-line effect occurring during the speech production process. In a following consonant context, an alveolar -t/d gesture is followed by a consonant gesture. For some words, the alveolar gesture will also be preceded by a consonant gesture. Deletion rates were greatest in interconsonantal environments, as has been previously observed (e.g., Wright, 1996; Zimmerer et al., 2014). During speech planning, articulatory simplification of consonant sequences can be accomplished by omitting the plan for an alveolar segment, especially in interconsonantal position. In a study employing electromagnetic articulography, Lichtman (2010) found that all speakers produced some -t words with no alveolar gesture, from which the absence of a planned alveolar gesture can be inferred. However, even planned segments may be susceptible to deletion during articulation, as proposed in frameworks like articulatory phonology (Browman & Goldstein, 1990, 1992). Successive consonants in codas and across word boundaries are prone to more gestural overlap than sequences in word onsets or word internally (e.g., Browman & Goldstein, 1992; Byrd, 1996; Marin & Pouplier, 2010). Gestural overlap in the extreme may result in the occlusion or attenuation of the -t/d gesture, leaving little or no acoustic trace (e.g., Browman & Goldstein, 1992, for English -t; Falahati & Mielke, 2011, for Persian final consonant clusters). The result is an acoustic deletion, an extreme type of acoustic reduction (see Ernestus, 2014). Deletions that occur during planning and during articulation will be indistinguishable in visual analysis of the speech signal alone.

In following vowel contexts, deletion during production requires a different explanation. Flapping context tokens did not show greater deletion with increasing FRC_{del}, but rather with increasing FRC_{flap}, essentially the contextual complement of FRC_{del}. Flapping context history thus predicted deletion in words in which flapping is likely, suggesting that for these words deletion was a consequence of the same process that creates flaps. Deletion in the flapping context can be modeled as extreme articulatory reduction (specifically, gestural undershoot or approximantization) of the flapping gesture (see Rhodes, 1992; Warner & Tucker, 2011).[5]

The differing processes evoked to explain on-line deletion phenomena in consonant and vowel contexts for -t/d parallels conclusions from a study of word-internal t/d deletion (Raymond et al., 2006) for t/d's in syllable onsets and codas. Intervocalic (onset) t/d tokens had high rates of flapping, with increased

https://doi.org/10.1017/S0954394516000041 Published online by Cambridge University Press

deletion rates in faster, more predictable, and more fluent speech, environments that encourage gestural approximantization. Deletions in tokens of t/d's in consonant clusters (coda tokens), on the other hand, were not as sensitive to these factors but rather were more common in environments interacting with planning processes (e.g., prominent syllables and less predictable phone sequences). Raymond et al. (2006) concluded that onset deletion resulted from approximantization of the flap articulation, as in the flapping environment of the current study, whereas coda deletion resulted from cluster simplification during planning or gestural overlap during production, as in consonant contexts of the current study.

Given the proposals for the on-line processes responsible for deletion, the effects of type-level cumulative contextual measures can be accommodated in usage-based models (e.g., Bybee, 2000, 2002) or representational models (e.g., Connine & Pinnow, 2006) of language. In these models, experiences with individual words or word variants in context are stored as abstract exemplars (or, in alternative models, as clusters of episodes; see, e.g., Goldinger, 1996; Johnson, 1997) that are weighted by frequency of occurrence. Exemplars contain acoustic and articulatory details of the tokens encountered by speakers. Evidence for frequency-weighted stored variants comes from studies of speech processing and speech production. For example, Pinnow and Connine (2014) found that subjects recognized high reduction rate schwa-reduced variants of words (e.g., *"corprate"* from *corporate*) better than low reduction rate schwa-reduced variants (e.g., *"crimnal"* from *criminal*). Pitt et al. (2011) found better recognition of the pronunciation variants of t/d words that were more frequently produced for a given word than of other noncanonical variants of those words (see also Ranbom & Connine [2007], for frequency-sensitive recognition of nasal flaps). In a production study, Bürki et al. (2010) showed that speakers of French produced schwa-reduced variants more quickly for words that had high estimated rates of schwa reduction than for words with low estimated rates of reduction.

Frequency-weighted representations thus predict processing accuracy and production facility, and their existence can also explain variable realization in production. In exemplar or usage-based models (Bybee, 2000, 2002; Pierrehumbert, 2001, 2006), production plans for words are based on their stored exemplars, which are based on experiences with words. The exemplars for words frequently encountered with the final segment deleted will be more numerous or stronger than those for words not encountered often with a deleted final segment. Subsequently, when saying words with higher deleted variant frequencies, speakers should more likely, or more readily, access speech plans that do not include the final segment than when saying words with lower deleted variant frequencies. Variant rates like those in the processing studies discussed were determined from phonetically annotated corpora (Connine, Ranbom, & Patterson, 2008), speaker estimations of variant frequencies (Bürki et al., 2010), or word variant production studies (Pitt et al., 2011). However, rates of word-final variants, like those for English -t/d words, can be estimated from context histories derived

https://doi.org/10.1017/S0954394516000041 Published online by Cambridge University Press

from a corpus of only orthographically transcribed speech. Estimation is possible because alternating (following the terminology of Timberlake, 1978), word-external phonetic environments are reflected in orthography.

In the case of -t/d deletion, words with final /t/ or /d/ with nonflapping preceding contexts that have higher FRC_{del} have occurred proportionally more often before orthographically consonant-initial words, an environment that encourages final segment deletions[6] and are thus more likely to have been experienced in deleted form. Speakers' representations of higher FRC_{del} words will presumably have been shaped by experience with these reduced forms to include more, or stronger, exemplars with -t/d deleted than in representations for words with lower FRC_{del}. Experience with deleted variants could then lead to lexemes for both deleted and nondeleted forms in the lexicon, consistent with proposals for representation of multiple variants (e.g., for categorical schwa reduction in English [Bürki & Gaskell, 2012], French [Bürki et al., 2010], and Dutch [Hanique et al., 2013], as well as nasal flapping of /nt/ in English [Ranbom & Connine, 2007]). Given multiple variant representations for each word, speakers would then more likely access speech plans that do not include the final /t/ or /d/ segment for higher FRC_{del} words than for lower FRC_{del} words.

As noted, FRC_{del} for word *w* is the predictability of a following consonant context given *w*. In speaking, context predictability could thus be used to project a phonological production plan that results in -t/d deletion even before the context word is planned. The result would be greater likelihood of -t/d deletion with increasing FRC_{del} in any context. Similarly words in flapping contexts with higher FRC_{flap} have occurred more often in a flapping environment and have more or stronger flapped exemplars in their representations (see Connine [2004], for evidence of the representation of flapped variants in frequently flapped words). These frequently flapped words are thus more likely to flap. High FRC_{flap} words will likely also possess more or stronger exemplars in which undershoot of the flap gesture has resulted in deletion, and so will be more likely to be produced with no -t/d as a consequence of experience as well.[7]

Although a cumulative type-level effect of one context type was found, there was no effect found of word frequency. The failure to find an effect of word frequency on -t/d deletion corroborates studies in which word frequency and an FRC measure of context history were considered together in statistical models of synchronic reduction (Raymond & Brown, 2012) and diachronic change (Brown & Raymond, 2012) in Spanish data. As noted, numerous earlier studies of various reduction phenomena had found a correlation between word frequency and rates of reduction or deletion, although some studies of the same phenomena did not. With regard to -t/d deletion, Walker (2012), for example, found no contribution of word frequency to predicting deletion in a corpus study after controlling lexical, phonological, and morphological factors. Importantly, earlier studies testing word frequency effects have not controlled words' reducing context frequencies (see Raymond & Brown [2012], for further discussion).

Phenomena not involving discrete phonetic variation at word boundaries have also been shown to be affected by contextual frequencies based on contextual

https://doi.org/10.1017/S0954394516000041 Published online by Cambridge University Press

variation. One such phenomenon is the durational shortening of words. Brown and Raymond (2014) found that a fast speech context encouraged durational reduction, as expected, but so also did the likelihood that a word occurs in the context of faster speech generally, a measure that can be called FRC_{rate}. Given the failure to find a word frequency effect on -t/d deletion in the current study or in other studies including contextual frequency measures, it might also be the case that previous demonstrations of the effect of word frequency on duration (e.g., Gahl, 2008) are better explained by FRC_{rate} effects, but the two variables (word frequency and FRC_{rate}) have not yet been included in the same model. In another study of acoustic durational shortening, Seyfarth (2014) found that a cumulative effect of a word's informativity from following lexical contexts (i.e., its average predictability from the following word across all contexts) could also predict durational shortening independent of its predictability in its production context. Because the cumulative effect of context rate and of informativity have not been tested together, it is similarly not known at this time whether the two factors have independent effects or not. Words that are more likely to occur in fast speech may also be susceptible to noncanonical segmental variation, because faster speech increases this type of variability (e.g., Fosler-Lussier & Morgan, 1999). A cumulative effect of speech rate, informativity, or other factors could explain lexical variability in the likelihood of schwa deletion, even in pairs of words that are structurally similar (e.g., high deletion likelihood *memory* vs. low deletion likelihood *mammary*; see Bybee, 2002). Finally, in a study of Spanish corpus data, Brown and Rivas (2012) found that a measure of FRC (specifically, the likelihood of a noun's use as sentential subject) predicted the likelihood of a reanalysis of nouns as subjects of the Spanish presentational (*haber* + noun phrase) construction, extending the usefulness of the FRC concept to morphosyntactic variation.

One question that arises given the effect of contextual history in the current study is whether the cumulative effect of context is at least in part a cumulative effect of a speaker's exposure to deleted forms. That is, FRC_{del} may simply be an index of the number of deletions a listener has encountered for each word, episodic traces of which lead to more deletion in production. This question cannot be answered without further study in which rates of occurrence in contexts encouraging deletion are compared to deletion rates. However, results from the processing literature already partially discussed suggest that there may be an effect of both variables. In addition to the effects of variant frequency on processing, Pinnow and Connine (2014) also found a transfer effect of form similarity in reduced-form processing in the absence of reduced-form exposure. Presumably the processing advantage conferred by similarity occurred based on speakers' knowledge of the acceptability of schwa deletion in the similar environments that licensed it (see also Mitterer & Ernestus, 2006; Pitt, 2009). Although the schwa-deleting environments are word internal, a study of Dutch -t word variants suggested that recognition of very reduced or deleted variants occurred after taking into account context of production (Janse, Nooteboom, & Quené, 2007). The processing results suggest that in production of -t/d words a

https://doi.org/10.1017/S0954394516000041 Published online by Cambridge University Press

higher predictability of the -t/d deleting environment could lead to a greater likelihood of deletion. This effect may be in addition to an effect of experience with deleted tokens in this environment. The two effects together would be compatible with hybrid processing models (Pierrehumbert, 2006; Sumner & Samuel, 2009), which recognize a need for both representational and inferential (Gaskell & Marslen-Wilson, 1998) processing of variants.

The finding that there is a cumulative effect of variable context on word variant production has other implications for theories of lexical representation. Lexically specific effects of context history suggest that a word's mental representation not only must include exemplars that reflect variant forms encountered in language use (as specified in the models of Bybee [2000, 2002], and Pierrehumbert [2001, 2006]), but also must encode distributional statistics about the contexts in which each variant has occurred. There is also evidence that contextual frequency information must extend to larger units and various levels of abstraction (see Arnon & Snider, 2010), and even to representation of social features, such as speaker gender, age, and accent (see Sumner, Kim, & McGowan, 2014). Studies have found effects of the probabilities of and predictability from surrounding words on word reduction, including -t/d deletion (e.g., Jurafsky et al., 2001a, 2001b), although these studies did not control contextual likelihoods. Regardless of which cumulative factors prove to be significant, there is a need for a word's representation to include a rich store of information about prior contexts' forms and frequencies.

Finally, there was no evidence in the data that a word's production in the deleting context increased deletion likelihood on its next production. Tamminga (2014) found evidence for form perseveration across successive occurrences of identical monomorphemic -t/d words. However, she did not examine the effect of prior production context or take into account the effect of cumulative exposure to deleting environments (FRC_{del}). Note also that in her corpus data, successive occurrences of a word were likely in the same discourse episode; the experimental paragraphs were, on the other hand, in distinct episodes on different topics, which has been shown to inhibit reduction resulting from repetition (Fowler, Levy, & Brown, 1997). Further study will be needed to assess effects of prior form, prior context, and discourse structure on the realization of words, as well as the time-course of consolidation of experience into exemplars or the acquisition and use of episodic traces of form experience on form production.

In conclusion, we have shown that, as in spontaneous speech, -t/d deletion is common in English read, narrative speech, and that -t/d deletion rates for words in this type of speech are, as in spontaneous speech, sensitive to a word's phonological structure and following phonological context. The effects of preceding and following segment contexts on -t/d deletion can be understood in terms of the articulatory processes of gestural overlap and undershoot during word production. However, there is, independently, a cumulative, type-level effect of experience with deleting contexts on -t/d deletion rates. The cumulative effect can be estimated by its frequency of occurrence in reducing contexts

https://doi.org/10.1017/S0954394516000041 Published online by Cambridge University Press

encouraging deletion (or FRC_{del}) using corpus distributions. The results add -t/d deletion to other phenomena for which experience with variable production contexts for word types has been shown to influence rates of variant form production in tokens. It is proposed that the predictive value of FRC measures, including FRC_{del} and FRC_{flap}, for variant production reflects the existence of multiple variant representations for words in the mental lexicon and leads to an account of variant lexeme origins through differential exposure to variable environments over time.

SUPPLEMENTARY MATERIAL

To view supplementary material for this article, please visit http://dx.doi.org/10.1017/S0954394516000041.

NOTES

1. One male speaker was bilingual in English (his predominant language now) and Spanish (his first language) from an early age (and had no discernable nonnative accent). Given that the task was conducted in monolingual (English) mode, any transfer effects from this single participant's Spanish were deemed to be minimal (see Grosjean, 1998).

2. A total of 18 different following consonants were used. The sequences /t/#/t/ and /d/#/d/ were not included to prevent ambiguities in identifying deletion. Words beginning with voiceless interdental fricatives were not included; however, *the* was used after both -t and -d (cf. Tagliamonte & Temple, 2005; Walker, 2012).

3. The ANOVA showed an effect of local context; that is, there was greater deletion when words were followed by a consonant than by a vowel in the paragraph (first or second) being read. However, there was no effect of repetition on deletion; thus, there was no evidence that presenting a word in the second paragraph that had also been read in the first paragraph increased the likelihood of its deletion relative to that for words shown only in the second paragraph. This lack of an effect of repetition held regardless of the contexts in which the word occurred in either paragraph.

4. Of the models examined, the final model had the lowest Akaike information criterion (AIC) and Bayesian information criterion (BIC), both provided in the output of *lmer()*. Confirmation that the final model gave the best fit with the variables under consideration was obtained by applying *drop1(),* which compares nested models using a likelihood ratio test; the output of this function indicated that no simpler model (i.e., a model in which individual terms or interactions have been removed) provided a better fit to the data (as measured by AIC and BIC).

5. There may be an explanation similar to the flapping account for deletion in glottalizing tokens; deletion would result from a weakening of the glottal gesture in words in a context favoring glottalization, which has been identified as a final /t/ preceded by a vowel, nasal, or liquid and followed by a stressed, consonant-initial word or a pause (see Eddington & Channer, 2010). An FRC_{glott} measure could be calculated for words using these criteria and would presumably predict not only glottalizations for these words, but, in cases of approximated glottal gestures, deletions as well.

6. Finer distinctions among following contexts could obviously be made, and separate FRC measures could be developed for different phone, or phone class, contexts.

7. Note, however, that some words may not have separate lexemes for deleted and nondeleted variants, consistent with evidence for gradient reduction phenomena in schwa deletion in some English schwa words (e.g., Bürki & Gaskell, 2012; Ranbom & Connine, 2007).

REFERENCES

Arnon, Inbal, & Snider, Neal. (2010). More than words: Frequency effects for multi-word phrases. *Journal of Memory and Language* 62:67–82.

Audacity Team. (2008). Audacity (version 1.2.3) [Computer program]. Available at: http://audacity.sourceforge.net/. Accessed July 2006.

https://doi.org/10.1017/S0954394516000041 Published online by Cambridge University Press

Baayen, R. Harald, Davidson, Donald J., & Bates, Douglas M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language* 59:390–412.
Bates, Douglas. (2005). Fitting linear mixed models in R. *R News* 5(1):27–30.
Bates, Douglas, Maechler, Martin, & Bolker, Benjamin. (2011). lme4: Linear mixed-effects models using S4 classes. R package version 0.999375-40. Available at: http://CRAN.R-project.org/package=lme4. Accessed November 2014.
Boersma, Paul, & Weenink, David. (2005). Praat: Doing phonetics by computer (version 5.2.18) [Computer program]. Available at: http://www.praat.org/. Accessed March 2011.
Browman, Catherine P., & Goldstein, Louis. (1990). Gestural specification using dynamically-defined articulatory structures. *Journal of Phonetics* 18:299–320.
_____. (1992). Articulatory phonology: An overview. *Phonetica* 49:155–180.
Brown, Esther L. (2004). *Reduction of syllable-initial /s/ in the Spanish of New Mexico and Southern Colorado: A usage-based approach*. Ph.D. dissertation, University of New Mexico.
Brown, Esther L., & Raymond, William D. (2012). How discourse context shapes the lexicon: Explaining the distribution of Spanish f-/h- words. *Diachronica* 29(2):139–161.
_____. (2014). Contextual frequency effects in Spanish phonology. Paper presented at the Georgetown University Round Table, Washington, D.C., March.
Brown, Esther L., & Rivas, Javier. (2012). Grammatical relation probability: How usage patterns shape analogy. *Language Variation and Change* 24:317–341.
Bürki, Audrey, Ernestus, Mirjam, & Frauenfelder, Ulrich H. (2010). Is there only one "fenêtre" in the production lexicon? On-line evidence on the nature of phonological representations of pronunciation variants in French schwa words. *Journal of Memory and Language* 62:421–437.
Bürki, Audrey, Ernestus, Mirjam, Gendrot, Cédric, Fougeron, Cécile, & Frauenfelder, Ulrich. (2011). What affects the presence versus absence of schwa and its duration: A corpus analysis of French connected speech. *Journal of the Acoustical Society of America* 130(6):3980–3991.
Bürki, Audrey, & Gaskell, M. Gareth. (2012). Lexical representation of schwa words: Two mackerels, but only one salami. *Journal of Experimental Psychology: Language, Memory, and Cognition* 38:617–631.
Bybee, Joan L. (2000). The phonology of the lexicon: Evidence from lexical diffusion. In M. Barlow & S. Kemmer (eds.), *Usage-based models of language*. Stanford: CSLI. 65–85.
_____. (2002). Word frequency and context of use in the lexical diffusion of phonetically conditioned sound change. *Language Variation and Change* 14:261–290.
Byrd, Dani. (1994). Relations of sex and dialect to reduction. *Speech Communication* 15:39–54.
_____. (1996). Influences on articulatory timing in consonant sequences. *Journal of Phonetics* 4:59–74.
Coetzee, Andries W. (2009). An integrated grammatical/non-grammatical model of phonological variation. In Y.-S. Kang, J.-Y. Yoon, H. Yoo, S.-W. Tang, Y.-S. Kang, Y. Jang, C. Kim, K.-A. Kim, & H.-K. Kang (eds.), *Current issues in linguistic interfaces*. Vol. 2. Seoul: Hankookmunhuasa. 267–294.
Connine, Cynthia M. (2004). It's not what you hear but how often you hear it: On the neglected role of phonological variant frequency in auditory word recognition. *Psychonomic Bulletin and Review* 11:1084–1089.
Connine, Cynthia M., & Pinnow, Eleni. (2006). Phonological variation in spoken word recognition: Episodes and abstractions. *Linguistic Review* 23:235–245.
Connine, Cynthia M., Ranbom, Larissa J., & Patterson, David J. (2008). On the representation of phonological variant frequency in spoken word recognition. *Perception & Psychophysics* 70:403–411.
Davies, Mark. (2008–). The Corpus of Contemporary American English: 450 million words, 1990–present. Available at: http://corpus.byu.edu/8. Accessed August 2012.
de Jong, Kenneth. (1998). Stress-related variation in the articulation of coda alveolar stops: Flapping revisited. *Journal of Phonetics* 26:283–310.
Dixon, Peter. (2008). Models of accuracy in repeated-measures design. *Journal of Memory and Language* 59:447–456.
Eddington, David, & Channer, Caitlin. (2010). American English has go? a lo? of glottal stops: Social diffusion and linguistic motivation. *American Speech* 85:338–351.
Ernestus, Mirjam. (2014). Acoustic reduction and the roles of abstraction and exemplars in speech processing. *Lingua* 142:27–41.
Ernestus, Mirjam, Lahey, Mybeth, Verhees, Femke, & Baayen, R. Harald. (2006). Lexical frequency and voice assimilation. *Journal of the Acoustic Society of America* 120:1040–1051.

https://doi.org/10.1017/S0954394516000041 Published online by Cambridge University Press

Falahati, Reza, & Mielke, Jeff. (2011). An ultrasound study of coronal stop deletion in Persian. *Canadian Acoustics/Acoustique Canadienne* 39:172–173.
Fasold, R. (1972). *Tense marking in Black English: A linguistic and social analysis*. Urban Language Series 8. Washington, D.C.: Center for Applied Linguistics.
File-Muriel, Richard J., & Brown, Earl K. (2011). The gradient nature of *s*-lenition in Caleño Spanish. *Language Variation and Change* 23:223–243.
Fosler-Lussier, Eric, & Morgan, Nelson. (1999). Effects of speaking rate and word frequency on conversational pronunciations. *Speech Communication* 29:1137–1158.
Fowler, Carol A., Levy, Elena T., & Brown, Julie M. (1997). Reductions of spoken words in certain discourse contexts. *Journal of Memory and Language* 37:24–40.
Gahl, Susanne. (2008). "Thyme" and "time" are not homophones: The effect of lemma frequency on word durations in spontaneous speech. *Language* 84:474–496.
Gaskell, M. Gareth, & Marslen-Wilson, William D. (1998). Mechanisms of phonological inference in speech perception. *Journal of Experimental Psychology: Human Perception and Performance* 24:380–396.
Godfrey, John J., Holliman, Edward C., & McDaniel, Jane. (1992). SWITCHBOARD: Telephone speech corpus for research and development. *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing* 1:517–520.
Goldinger, Stephen D. (1996). Words and voices: Episodic traces in spoken word identification and recognition memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 22:1166–1183.
Grosjean, François. (1998). Studying bilinguals: Methodological and conceptual issues. *Bilingualism: Language and Cognition* 1:131–149.
Guy, Gregory R. (1980). Variation in the group and the individual: The case of final stop deletion. In W. Labov (ed.), *Locating language in time and space*. Academic Press: New York. 1–36.
Guy, Gregory R., & Boberg, Charles. (1997). Inherent variability and the obligatory contour principle. *Language Variation and Change* 9:149–164.
Guy, Gregory R., Hay, Jennifer, & Walker, Abby. (2008). Phonological, lexical, and frequency factors in coronal stop deletion in early New Zealand English. Poster presented at The 11th Conference on Laboratory Phonology, Wellington, July.
Hanique, Iris, Ernestus, Mirjam, & Schuppler, Barbara. (2013). Informal speech processes can be categorical in nature, even if they affect many different words. *Journal of the Acoustical Society of America* 133:1644–1655.
Janse, Esther, Nooteboom, Sieb G., & Quené, Hugo. (2007). Coping with gradient forms of /t/-deletion and lexical ambiguity in spoken word recognition. *Language and Cognitive Processes* 22:161–200.
Johnson, Keith. (1997). Speech perception without speaker normalization: An exemplar model. In K. Johnson, & J. W. Mullennix (eds.), *Talker variability in speech processing*. San Diego: Academic. 145–165.
Jurafsky, Daniel, Bell, Alan, Fosler-Lussier, Eric, Girand, Cynthia, & Raymond, William D. (1998). Reduction of English function words in Switchboard. *Proceedings of the International Conference on Spoken Language Processing* 7:3111–3114.
Jurafsky, Daniel, Bell, Alan, Gregory, Michelle, & Raymond, William D. (2001a). The effect of language model probability on pronunciation reduction. *Proceedings of ICASSP-01* 2:801–804.
_____. (2001b). Probabilistic relations between words: Evidence from reduction in lexical production. In J. Bybee, & P. Hopper (eds.), *Frequency and the emergence of linguistic structure*. Amsterdam: John Benjamins. 229–254.
Kliegl, Reinhold, Wei, Ping, Dambacher, Michael, Yan, Ming, & Zhou, Xiaolin. (2011). Experimental effects and individual differences in linear mixed models: Estimating the relationship between spatial, object, and attraction effects in visual attention. *Frontiers in Psychology* 1:238.
Labov, William. (1972). *Sociolinguistic patterns*. Philadelphia: University of Pennsylvania Press.
Lichtman, Karen. (2010). *Testing articulatory phonology: Variation in gestures for coda /t/*. Paper presented at ILLS (Illinois Language and Linguistic Society) 2, University of Illinois at Urbana-Champagne, May.
Marin, Stefania, & Pouplier, Marianne. (2010). Temporal organization of complex onsets and codas in American English: Testing the predictions of a gestural coupling model. *Motor Control* 14:380–407.
Mitterer, Holger, & Ernestus, Mirjam. (2006). Listeners recover /t/'s that speakers reduce: Evidence from /t/-lenition in Dutch. *Journal of Phonetics* 34:73–103.

https://doi.org/10.1017/S0954394516000041 Published online by Cambridge University Press

Neu, Helene. (1980). Ranking of constraints on American /t,d/ deletion in American English: A statistical analysis. In W. Labov (ed.), *Locating language in time and space*. Academic Press: New York. 37–54.
Oshika, Beatrice T., Zue, Victor W., Weeks, Rollin V., Neu, Helene, & Auerbach, Joseph. (1975). The role of phonological rules in speech understanding research. *IEEE Transactions on Acoustics, Speech, and Signal Processing* ASSP-23:104–112.
Patrick, Peter L. (1991). Creoles at the intersection of variable processes: -t, -d deletion and past-tense marking in the Jamaican mesolect. *Language Variation and Change* 32:171–189.
Phillips, Betty S. (2006). *Word frequency and lexical diffusion*. New York: Palgrave Macmillan.
Pierrehumbert., Janet B. (2001). Exemplar dynamics: Word frequency, lenition, and contrast. In J. Bybee & P. Hopper (eds.), *Frequency and the emergence of linguistic structure*. Amsterdam: John Benjamins. 137–157.
_____. (2006). The statistical basis of unnatural alternation. In L. Goldstein, D. H. Whalen, & C. Best (eds.), *Laboratory phonology VIII, varieties of phonological competence*. Berlin, Germany: Mouton de Gruyter. 81–107.
Pinnow, Eleni, & Connine, Cynthia M. (2014). Phonological variant recognition: Representation and rules. *Language and Speech* 57:42–67.
Pitt, Mark. (2009). How are pronunciation variants of spoken words recognized? A test of generalization to newly learned words. *Journal of Memory and Language* 61:19–36.
Pitt, Mark, Dilley, Laura, Johnson, Keith, Kiesling, Scott, Raymond, William, Hume, Elizabeth, & Fosler-Lusier, Eric. (2007). Buckeye corpus of conversational speech. 2nd release. Columbus: Department of Psychology, Ohio State University. Available at: www.buckeyecorpus.edu. Accessed November 2003.
Pitt, Mark, Dilley, Laura, & Tat, Michael. (2011). Exploring the role of production frequency in recognizing pronunciation variants. *Journal of Phonetics* 39:304–311.
R Development Core Team. (2010). R: A language and environment for statistical computing. Vienna: R Foundation for Statistical Computing. Available at: http://www.R-project.org/. Accessed April 2011.
Ranbom, Larissa J., & Connine, Cynthia M. (2007). Lexical representation of phonological variation in spoken word recognition. *Journal of Memory and Language* 57:278–298.
Raymond, William D., & Brown, Esther L. (2012). Are effects of word frequency effects of context of use? An analysis of initial fricative reduction in Spanish. In S. Th. Gries & D. S. Divjak (eds.), *Frequency effects in language*. Vol. 2: *Learning and processing*. The Hague: Mouton de Gruyter. 35–52.
Raymond, William D., Dautricourt, Robin, & Hume, Elizabeth (2006). Word-medial /t-d/ deletion in spontaneous speech: Modeling the effects of extra-linguistic, lexical, and phonological factors. *Language Variation and Change* 18:55–97.
Rhodes, Richard A. (1992). Flapping in American English. In W. U. Dressler, M. Prinzhorn, & J. R. Dennison (eds.), *Proceedings of the 7th International Phonology Meeting*. Turin: Rosenberg and Sellier. 217–232.
Santa Ana, Otto. (1991). *Phonetic simplification processes in the English of the Barrio: A cross-generational study of the Chicanos of Los Angeles*. Ph.D. dissertation, University of Pennsylvania. Retrieved from ProQuest Dissertations and Theses (Access Order No. 9200383).
Seyfarth, Scott. (2014). Word informativity influences acoustic duration: Effects of contextual predictability on lexical representation. *Cognition* 133:140–155.
Sumner, Meghan, Kim, Seung Kyung, & McGowan, Kevin B. (2014). The socially weighted encoding of spoken words: A dual route approach to speech perception. *Frontiers in Psychology* 4:1015.
Sumner, Meghan, & Samuel, Arthur G. (2009). The effect of experience on the perception and representation of dialect variants. *Journal of Memory and Language* 60:487–501.
Tagliamonte, Sali, & Temple, Rosalind. (2005). New perspectives on an ol' variable: (t,d) in British English. *Language Variation and Change* 17:281–302.
Tamminga, Meredith J. (2014). *Persistence in the production of linguistic variation*. Ph.D. dissertation, University of Pennsylvania.
Timberlake, Alan. (1978). Uniform and alternating environments in phonological change. *Folia Slavica* 2:312–328.
Torreira, Francisco, & Ernestus, Mirjam. (2011). Vowel elision in casual French: The case of vowel /e/ in *c'était*. *Journal of Phonetics* 39:50–58.
Van Heuven, Walter J. B., Mandera, Paweł, Keuleers, Emmanuel, & Brysbaert, Mark. (2014). Subtlex-UK: A new and improved word frequency database for British English. *Quarterly Journal of Experimental Psychology* 67:1176–1190.

https://doi.org/10.1017/S0954394516000041 Published online by Cambridge University Press

Walker, James A. (2012). Form, function, and frequency in phonological variation. *Language Variation and Change* 24:397–415.
Warner, Natasha, & Tucker, Benjamin V. (2011). Phonetic variability of stops and flaps in spontaneous and careful speech. *Journal of the Acoustical Society of America* 130:1606–1617.
Wolfram, Walt. (1969). *A sociolinguistic description of Detroit Negro speech*. Washington, D.C.: Center for Applied Linguistics.
Wright, Richard. (1996). *Consonant clusters and cue preservation in Tsou*. Ph.D. dissertation, University of California, Los Angeles.
Zimmerer, Frank, Scharinger, Mathias, & Reetz, Henning. (2011). When BEAT becomes HOUSE: Factors of word final /t/-deletion in German. *Speech Communication* 53:941–954.
_____. (2014). Phonological and morphological constraints on German /t/-deletions. *Journal of Phonetics* 45:64–75.
Zue, Victor W., & Laferriere, Martha. (1979). Acoustic study of internal /t,d/ in American English. *Journal of the Acoustic Society of America* 66:1039–1050.

https://doi.org/10.1017/S0954394516000041 Published online by Cambridge University Press

APPENDIX A

Experimental materials

Table A1 shows the words used in the experiment in each of the three lists, with a range of matched features among words in each list. Preceding the table, an example stimulus paragraph is provided (P12V). The target words are presented in capital letters in the example paragraph, but were not specially marked in the materials presented to participants.

> My mother BOUGHT enough plane tickets for my family to take a trip to California. The tickets were a GIFT after we saw a terrible INCIDENT occur really close to home; near our APARTMENT is a YARD in which a woman DIED after she was unable to REACT in time to AVOID a falling tree limb during a storm. We were all pretty shaken up. My mom was feeling GUILT about our witnessing everything, so we were to go on this trip as a way to LIFT up our spirits. Once we were all SEATED on BOARD expecting to leave, however, the captain says, "We have been DELAYED an hour because of weather." I was speechless! We were actually stuck for many, many hours waiting for the storm to pass. My young brothers were so TIRED in the plane because of the delay they began to ACT a little crazy. I ADMIT I even eventually WENT a little nuts sitting on the tarmac! We all WANTED our plane to take off anyway—even with the lightning. However the crew knew the BEST approach was to delay the departure. We will surely FAULT all the rain for ruining the beginning of our vacation!

https://doi.org/10.1017/S0954394516000041 Published online by Cambridge University Press

https://doi.org/10.1017/S0954394516000041 Published online by Cambridge University Press

TABLE A1. *Words comprising the three word lists used in the experiment, together with each word's controlled features, compared across lists*

−t/-d	Word class	Previous phone	List 1	Log_{10} freq	FRC_{del}	List 2	Log_{10}freq	FRC_{del}	List 3	Log_{10} freq	FRC_{del}
−t/-d	N	/l/	Guilt	2.89	.280	Fault	3.73	.268	Gold	3.61	.592
−d	N	/r/	Board	3.51	.391	Yard	3.11	.317	Card	3.64	.446
−t	N_{-ment}	/n/	Incident	2.95	.370	Apartment	3.63	.416	Engagement	2.82	.441
−d	$V_{-əd}$	/V,n,s/	Wanted	4.41	.864	Seated	2.64	.300	Wasted	2.93	.490
−d	V_{-d}	V	Died	3.90	.297	Delayed	2.44	.505	Said	4.75	.481
−t	V_{-t}	/V,n,s/	Bought	3.64	.541	Went	4.32	.588	Crossed	2.94	.654
−t/-d	$V_{pres.t}$	V	Admit	3.48	.643	Avoid	3.09	.681	Regret	3.14	.518
−t	N or V	/k/	Act	3.75	.314	React	2.59	.629	Fact	3.94	.382
−t	N or V	/f/	Gift	3.52	.542	Lift	3.24	.607	Craft	2.46	.378
−t/-d	Adj	/s,r/	Tired	3.76	.102	Best	4.34	.665	Fast	3.85	.352
Average				3.58	.433		3.31	.497		3.41	.473

APPENDIX B

*Complete results for GLMM final model predicting deletion, including nonsignificant predictors (*n = *1713)*

Parameters	*b*	SE	*p*	Sig.	*n*	% deletion
(Intercept)	−2.70	.86	.001764	**		
Following context (vs. *consonant*)					851	45
Vowel	–.65	.43	.136934		862	30
Preceding context (vs. *nonflapping*)					754	49
Flapping	1.46	.98	.137055		959	28
FRC_{del}	6.67	1.70	8.5*e*–05	***	Continuous	Variable
Syllable stress (vs. *stressed*)					1370	36
Unstressed	−1.29	.98	.186293		343	43
Gender (vs. *male*)					546	50
Female	.50	.49	.308070		1167	31
Following context:preceding context	.93	.25	.000192	***	Interaction	
Following context:FRC_{del}	−1.94	.77	.011623	*	Interaction	
Following context:syllable stress	1.39	.29	2.0*e*–06	***	Interaction	
Following context:gender	–.14	.32	.657444		Interaction	
Preceding context:FRC_{del}	−7.03	1.98	.000389	***	Interaction	
Preceding context:syllable stress	–.67	.70	.334365		Interaction	
Preceding context:gender	.75	.26	.003208	**	Interaction	
FRC_{del}:syllable stress	2.40	2.07	.247054		Interaction	
FRC_{del}:gender	–.31	.78	.691830		Interaction	
Syllable stress:gender	1.03	.31	.000859	***	Interaction	

Note: *** $p < .001$; ** $p < .01$; * $p < .05$.

https://doi.org/10.1017/S0954394516000041 Published online by Cambridge University Press