JOURNAL ARTICLE

The link between syllabic nasals and glottal stops in American English

Lisa Davidson, Shmico Orosco and Sheng-Fu Wang
Department of Linguistics, New York University, New York, NY, US
Corresponding author: Lisa Davidson (lisa.davidson@nyu.edu)

Examples of syllabic nasals in English abound in phonological research (e.g., Hammond, 1999; Harris, 1994; Wells, 1995), but there is little explicit discussion about the surrounding consonant environments that condition syllabic nasals. This study examines the production of potential word-final syllabic nasals in American English following preceding consonants including oral stops, glottal stops, fricatives, flap, and laterals. The data come from a laboratory study of read speech with speakers from New York and other regions, a corpus of read speech with speakers from the Pacific Northwest and Northern Cities, and a spontaneous telephone speech corpus. Acoustic analysis indicates that [ŋ̩] is only prevalent after [ʔ], with some extension to [d] or [ɾ]. Variation in rates of [ŋ̩] versus [ən] is found across the speakers in a group, not within individual speakers. An articulatory sketch to account for the prevalence of [n̩] after coronal and glottal stops is laid out. To link this realization to the presence of the [ʔ] allophone in pre-syllabic nasal environment, previous analyses of acoustic enhancement proposed for glottally-reinforced [tˀ] in coda position (e.g., Keyser & Stevens, 2006) are extended to the syllabic nasal case.

Keywords: Glottal stops; syllabic nasals; flaps; glottal reinforcement; connected speech

1. Introduction

Accounts of the appearance of syllabic nasals in English for both British (especially Received Pronunciation) and American varieties have tended to be based on impressionistic or introspective data. In (1), potential contexts for syllabic nasals in word-final position are illustrated using examples from these papers. As these examples show, almost all of the consulted sources mention the post-/t/ or glottal stop environment as a possible place for a final syllabic nasal, but many other possible environments have also been asserted. For example, Simpson (2005) suggests that there might be lexical idiosyncrasy for some words when [k] precedes a potential syllabic nasal; for example, *taken* could surface either as [teɪkən] or [teɪkn̩] (see also Szigetvári, 2002 for potential assimilation of [n̩] after a non-coronal stop). Wells (1995), on the other hand, explicitly states that the probability of a word-final syllabic nasal following a sonorant consonant, in words like *sullen* or *common*, is approaching zero. Several of the sources cited in (1) also mention other potential environments where a syllabic nasal could be found, including before a word-final [t] or [s], as in *absent, accident, residence*, stem-finally but preceding a morpheme, as in *threatening, reasonable*, or even word-internally in a monomorpheme, as in *lavender, ordinary* (Cruttenden, 2008; Hammond, 1999; Heselwood, 2007; Mora Bonilla, 2003). Such cases are not included in the laboratory and corpus studies presented in this paper, and will not be considered further.

(1)   a.   Following [ʔ] (most American English) or [t] (some British English):
             button, frighten
             (Carley, Mees, & Collins, 2017; Cruttenden, 2008; Hammond, 1999;
             Harris, 1994; Heselwood, 2007; Keyser & Stevens, 2006; Mora Bonilla,
             2003; Shockey, 2003; Simpson, 2005; Trager & Bloch, 1941; Wells, 1995)

       b.   Following [ɾ] or [d]: *sudden, hidden*
             (Carley et al., 2017; Heselwood, 2007; Keyser & Stevens, 2006; Mora
             Bonilla, 2003; Wells, 1995)

       c.   Following some fricatives: *seven, brazen*
             (Carley et al., 2017; Hammond, 1999; Keyser & Stevens, 2006; Mora
             Bonilla, 2003; Wells, 1995)

       d.   Following non-coronal stops: *happen, reckon*
             (Polgárdi, 2014; Rubach, 1996; Szigetvári, 2002)

The references in (1), many of which discuss British English, suggest that [n̩] should be most likely to be found following a coronal consonant. Yet, there is very little explicit discussion of the relationship between the preceding consonant and the likelihood of the production of a syllabic nasal. One exception is the Roach (2009) textbook, where he (anecdotally) observes that in British English, [n̩] is most likely after alveolar stops and fricatives (but not postalveolar affricates), rare after velar stops, and variable after bilabial stops. He also remarks that [n̩] is more common than [ən] after [f] and [v]. Moreover, in the case of American English, what is produced as [t] in many varieties of British English is said to be produced as either a true glottal stop or as a glottally-reinforced [tʔ] (e.g., Byrd, 1994; Eddington & Channer, 2010; Huffman, 2005; Kahn, 1980; Keyser & Stevens, 2006; Pierrehumbert, 1995; Roberts, 2006; Seyfarth & Garellek, 2015, 2020). Since the accounts referenced in (1) are mostly introspective, the study in this paper aims to clarify the distribution of word-final syllabic nasals in American English, in order to determine whether they are particularly widespread, or instead limited to occurring after specific sounds.

   The main focus of many previous accounts is instead on whether potential syllabic nasals are derived from /ən/, or whether the syllabic nasal itself is underlyingly represented. Many authors assume that /ən/ is the underlying sequence (Gussman, 1991; Harris, 1994; Polgárdi, 2014; Shockey, 2003; Toft, 2002). Trager and Bloch (1941, p. 232) state that "there is often free (stylistically determined) variation" between the syllabic and vowel variants in American English, and Rubach (1996) notes that in British RP, the presence of schwa is optional, but syllabic nasals are mostly obligatory in American English. Stevens and Keyser (2010) state that syllabic nasals are due to overlap between the vocalic and nasal consonant articulations, and assume that the resultant syllabic nasal should contain acoustic cues to the existence of the vowel, such as a longer duration and low-frequency amplitude that is more attributable to a vowel than to a nasal consonant by itself. Wells (1995) stipulates that the schwa is present in the UR, in part because he notes that impressionistically the presence of schwa is variable for some speakers or dialects, and he proposes a rule that converts a schwa + sonorant C sequence to a syllabic consonant. Such a rule is applied with variable frequency, being most likely after a stressed syllable ending in an alveolar plosive (e.g., *button, sudden*) and least likely after a sonorant consonant (e.g., *sullen, common*).

   Mora Bonilla (2003) rejects accounts with an underlying vowel, instead positing that there is no underlying vowel, but if a nasal consonant follows a lower sonority sound like a stop or a fricative, then it will be forced to assume syllabicity to improve the sonority profile of the word (e.g., *mitten:* /mɪtn/ → [mɪtn̩]). Mora Bonilla observes that a limitation of his account is that at least according to the examples that some phonologists have

previously given, words like *foreign*, which is claimed to have a syllabic nasal [fɒɹn̩] in Southern British English, should not occur because it does not contain the proper sonority environment for syllabic consonant formation.

Whether or not examples like *foreign* really pose a problem for Mora Bonilla's account is an empirical question given the dearth of instrumental studies. An early study to examine the proportion of potential syllabic nasals in speech, using a spoken speech corpus, is Roach et al. (1992). This study examines both American English and British English using the syllabic nasal and [ən] notations provided by the transcribers in the TIMIT corpus and the British SCRIBE corpus. For both corpora, they compare the numbers of words transcribed with [n̩] to those that have a potential syllabic nasal environment but are transcribed with [ən]. Although Roach et al. do not specify whether the examples are both word-internal (e.g., *lightening* [laɪtn̩ɪŋ]) and word-final (e.g., *button*), it is assumed that they are since both cases are previously discussed in the paper. Results show that in American English, 9.3% of potential [n̩] (total potential N = 7135) were transcribed as syllabic, while 8.9% of potential [n̩] (potential N = 2216) in British English were transcribed as syllabic. While the phonetic criteria used by the transcribers are not provided, it is nevertheless noteworthy that these proportions are quite small.

Toft (2002) reports on a small production study in British English in which the preceding consonantal environments before potential [n̩#] are divided between [p], [t], and [k]. Toft also provides some spectrographic illustrations of her criteria for determining the presence of a schwa. Data from 8 participants producing a total of 324 potential [n̩#] words shows that [n̩] is produced 27% and 30% of the time after [p] and [k] respectively, and 85% of the time after [t]. Using just a small set of preceding contexts, these results suggest that preceding coronal consonants are more likely to condition [n̩#]. For American English, Eddington and Savage (2012) examine the [ʔ(ə)n] environment specifically and show that the presence of the schwa may be dialectally conditioned. In particular, young, female Utahans are most likely to produce [ʔən], followed by the youngest male Utahans in their sample, as compared to non-Utahans who almost exclusively produce [ʔn̩].

Taken together, the few available phonetic studies suggest that it is likely that [n̩] occurs after a coronal consonant, and in particular a coronal stop, but that syllabic nasals may otherwise be rarer than is assumed in some of the older impressionistic literature. Roach (2009) notes that the likelihood of syllabic [n̩] after coronal stops is because 'nasal release' occurs in this environment, where the tongue tip stays raised for the coronal stop that is oral for [t] or [d] and becomes nasal for a subsequent [n]. Instead of producing a schwa, then, the [n] is produced syllabically (see a similar description in Hall, 2006; Ladefoged & Johnson, 2014; Zue & Laferriere, 1979). However, as shown in (1), plenty of sources do assume that syllabic nasals can and do occur after non-coronal and non-stop consonants as well.

1.1. Research questions

The goal of this paper is two-fold. First, we present data from several sources to examine how dialectal background and, to a limited degree, speech style affects the distribution of syllabic nasals in word-final position in American English. The first study is conducted in a lab setting, with speakers from the New York area and from other regions reading sentences containing words with potential word-final syllabic nasals. The second study is a similar analysis of the appropriate words identified in the University of Washington/Northwestern University (UW/NU) Corpus (Panfili, Haywood, McCloy, Souza, & Wright, 2017) in order to investigate whether speakers from other dialect regions in the United States, namely the Pacific Northwest and Northern Cities, have a similar or different pattern of results. These studies are supplemented with data from spontaneous speech from the Fisher Conversational Telephone Speech (CTS) corpus

(Cieri, Graff, Kimball, Miller, & Walker, 2004), in order to examine whether the rates of syllabic nasals found for spontaneous speech are similar to those in read speech. This corpus spans the North, Midland, South, and West of the United States.

The second goal of this research is to examine whether certain phonetic contexts condition the particular realization of [ən]/[n̩] word-finally in American English by varying the preceding consonant contexts, including non-coronal and coronal oral stops, glottal stops, fricatives, and laterals. Based on the discussions from previous research, we hypothesize that the presence of a schwa is conditioned by the preceding phoneme, in particular, coronal stops. This would suggest an articulatory motivation for the distribution of syllabic nasals along the lines sketched by Roach (2009). That is, the continued tongue tip raising from the coronal oral stop to the nasal consonant blocks the acoustic realization of the vowel and is produced on the surface as a syllabic nasal instead. We also examine whether there is lengthened duration of the nasal consonant, which could be consistent with Stevens and Keyser's (2010) suggestion that there may be overlap of the schwa and nasal consonant articulations.

Interpreting an outcome consistent with an effect of a preceding coronal stop is complicated by the issue of how the surface variants of coronal stops in American English interact with the realization of either [ən] or [n̩]. In the types of disyllabic target words included in this study, /t/ typically surfaces as [ʔ] (for most words, the relationship with /t/ is established by paradigmatically related words, e.g., ro[t]/ro[ʔ]en, swee[t]/swee[ʔ]en, for example), and /d/ may surface as either [d] or [ɾ] (e.g., woo[d]/woo[ɾ/d]en, hi[d]/hi[ɾ/d]en). To be more specific, the symbol [ʔ] is used in this paper as a stand-in for multiple phonetic realizations, including a glottally reinforced [tʔ], a full glottal stop, and a period of glottalization (creaky voice) with no closure (Huffman, 2005; Pierrehumbert, 1995; Seyfarth & Garellek, 2020). The specific implementations will be discussed in more detail throughout the paper.

If the coronal stop that might articulatorily condition [n̩] is actually a glottal stop or a period of glottalization, can an articulatory pressure for [n̩] still be determined? To preview the results of the study, syllabic nasals are indeed especially likely after /t/, which is always produced as one of these glottalized variants. Therefore, in the discussion, we develop a possible explanation for why American English might have developed [ʔ] in precisely this environment, and what the articulatory relationship between [ʔ], [d/ɾ], and a [n̩] realization could be.

2. Production Data
2.1. Laboratory study
2.1.1. Participants
The participants were 25 monolingual speakers of American English (ages 19–32, 3M, 22F). All of the speakers had been living in the New York metropolitan area for at least a year, with 15 of the speakers having always lived in New York City or suburban New Jersey, Connecticut, or Long Island. Of the remaining speakers, the home states included California (3), Georgia (1), Illinois (2), Missouri (1), Oregon (1), Vermont (1), and Virginia (1). Participants were paid $5 for their time. This research was approved by the Institutional Review Board at New York University.

2.1.2. Stimuli and Procedure
The stimuli for this study included 40 target word-final [ən]/[n̩] words.[1] Speakers also produced 10 words containing [əm]/[m̩] endings, but these were not included in the current study. The [ən]/[n̩] words were evenly divided among preceding surface

---

[1] The label [ən]/[n̩] is used to indicate a potential environment for the realization of a syllabic nasal, and to remain agnostic at this stage about what the most likely implementation is.

consonant types as follows, with 8 words per preceding consonant type: [-coronal] stops, fricatives, [l], [d/ɾ], and [ʔ]. The words in the [ʔ] category are mostly part of a paradigm where [t] is realized word-finally in a related form (e.g., *rotten* and *ro[t]*, *eaten* and *ea[t]*). These words are all spelled with 't' or 'tt.' All of the speakers in this study pronounced these words with either a true glottal stop or a period of glottalization preceding the [ən]/[n̩] portion, and never with [t], which is why this category is referred to as [ʔ]. The words in the [d/ɾ] category were sometimes pronounced as [d] and other times as [ɾ]. The relationship between the specific production of /d/ as either [d] or [ɾ] and [ən] or [n̩] is addressed in the analysis. The complete list of stimuli is shown in **Table 1**.

Because of the relative sparsity of words with final [ən]/[n̩], frequency could not be easily controlled in developing this wordlist. However, frequency will be included as a linear factor where relevant in the analysis. It was also difficult to balance the number of morphologically simple versus complex words, since many words with word-final [ən]/[n̩] are morphologically complex, so we have not tried to control that factor in the stimuli.

The words in **Table 1** were combined into 26 sentences in phrase medial position, in sentences that were intended to avoid intonational boundaries or other pauses after the potential syllabic nasal. Example sentences are given in (2), and the whole list is in the Appendix. A total of 989 words were analyzed; 11 were removed because of mispronunciation, hesitations, or pauses right after the potential syllabic nasal.

Example sentences

    a.   We had to **loosen** the wheel on the **wagon** to remove it.
    b.   Miguel's symptoms may **worsen** if the doctors **weaken** his dosage.
    c.   Gail has never **eaten** at Burger **Heaven** in Brooklyn.

For the recording, participants were seated in a sound-treated room and were given a randomized list of the 26 sentences. They were told to read each sentence at a rate that felt natural to them. They read each sentence once aloud into a Shure SM10A headworn microphone attached to a Tascam DR-40 digital recorder. The sentences were recorded to uncompressed wav files at 44kHz. The recording session took about seven minutes to complete.

2.1.3. Data Analysis

Textgrids in Praat (Boersma & Weenink, 2020) were created for each file. For each target word, the following information was marked: whether or not a schwa was present, the intervals corresponding to the nasal and the schwa (if present), whether [d/ɾ] was produced as [ɾ] or as [d], and whether [ʔ] was produced as an actual stop or as a period of glottalization. All determinations were made relying on both the spectrogram and on perceptual information. For the [d/ɾ] case, [d] was marked if there was a period of

Table 1: Words used in the production study.

[-cor] stops	fricatives	[l]	[ʔ]	[d/ɾ]
ripen	loosen	stolen	button	hidden
deepen	worsen	sullen	beaten	wooden
sicken	risen	woolen	written	redden
weaken	chosen	fallen	eaten	sudden
urban	deafen	swollen	straighten	trodden
ribbon	stiffen	Allen	rotten	widen
dragon	heaven	Helen	tighten	leaden
wagon	proven	melon	sweeten	broaden

closure (with no formants or acoustic information other than voicing during the closure) followed by an observable burst. For [ɾ], there could be a short period of closure, but no visible burst, or there could be weakening to an approximant that was produced with lower intensity formants (consistent with the phonetic implementations found in Fukaya & Byrd, 2005; Herd, Jongman, & Sereno, 2010; Warner, Fountain, & Tucker, 2009; Zue & Laferriere, 1979). All instances of [ʔ] had at least two pulses of glottalization on the preceding vowel and sometimes on the following nasal, but to be marked as a true glottal stop, there had to be at least 20 ms of closure preceding the nasal or the schwa, if one was present. Otherwise, the glottal stop was realized as a period of glottalization with no oral closure (see **Figure 6** in Section 3.3).

Schwas were marked by looking for changes in intensity compared to adjacent sounds and differences in formant structure between the vowel and the following nasal. There was typically a very obvious intensity difference between the schwa and the following nasal, with the amplitude of the higher frequencies in the nasal (corresponding to frequencies higher than F3 in the vowel) lowering dramatically. Likewise, the nasal consonant typically had a nasal formant that was concentrated at a different frequency than F2 of the preceding vowel. The boundary between vowel and nasal was very clear and easy to demarcate in nearly all of the cases. The only potentially difficult case concerned preceding [l]. In this environment, a schwa was marked present when F2 rose and F3 lowered after the characteristic lowering of F2 and raising of F3 that characterizes intervocalic /l/ in American English (e.g., Recasens, 2012). These changes in formants were usually also accompanied by an increase in intensity between from the [l] to the schwa. In the very few cases that were difficult to determine, the first two authors together decided whether a vowel was present. Several examples of presence and absence of a schwa, as well as variation in [d/ɾ], and [ʔ] are shown in **Figure 1**.

The labeling and segmenting procedure was carried out initially by the second author, who received training in acoustic phonetics and speech science classes, and further training specific to the segments being examined in this study. The first author then reviewed all of the Praat textgrids, and in cases where there was disagreement (about 15% of the data), the first and second author then met together to resolve what the proper labels and segment boundaries should be. After this consensus procedure, no further data were removed.

2.1.4. Results

For the analysis of the presence versus absence of schwa, results were analyzed using a mixed effects logistic regression using *lme4* in R (Bates et al., 2018). The binomial dependent variable was the production of either a syllabic nasal or a schwa, with preceding consonant category ([-cor] stop, fricative, [l], [d/ɾ], [ʔ]) as a fixed effect. The preceding consonant category was sum coded, with the level 'fricative' held out. In addition, in order to further investigate whether the New York speakers exhibit any group differences compared to the remaining speakers from more diverse places, a factor of region (NY, other) is also a fixed effect and was sum-coded. The model also included the interaction of preceding consonant category and region. Speaker and word are included in the model as random intercepts.

For the analysis of the duration of the nasal in syllabic versus [ə] contexts, a linear mixed effects regression was carried out using *lme4* and *lmerTest* (Kuznetsova, Brockhoff, & Christensen, 2013). The dependent variable was the duration of the nasal, and the fixed factors are preceding consonant category ([-cor] stop, fricative, [l], [d/ɾ], [ʔ]) and schwa presence ([ən], [n̩]), and their interaction. Speaker and word are included in the model as random intercepts.

The proportions of syllabic nasals versus schwas are shown in **Figure 2**. Results given in **Table 2** indicate that preceding consonant context is significant for all of the contexts, but this only means that all of the values for these contexts are either significantly above or below the mean for the rate of schwa presence over all of the contexts. More critically, results in **Table 3** from Tukey tests using the *multcomp* package in R (Hothorn, Bretz, & Westfall, 2008) indicate significant differences between [d/ɾ] and all other contexts, and between [ʔ] and all other contexts. There are no significant differences between [-cor] stops, fricatives, or [l]. These results show that speakers nearly always produce the schwa variant in these contexts, are most likely to produce [n̩] following glottalization, and have intermediate rates of [n̩] following [d/ɾ].

As for speaker region, there is no significant main effect, but there is an interaction for non-New York speakers with [ʔ]. The significant interaction for [ʔ] is a result of the especially big difference between the rate of syllabic nasals produced by New Yorkers (63%) as compared to the other speakers (91%).
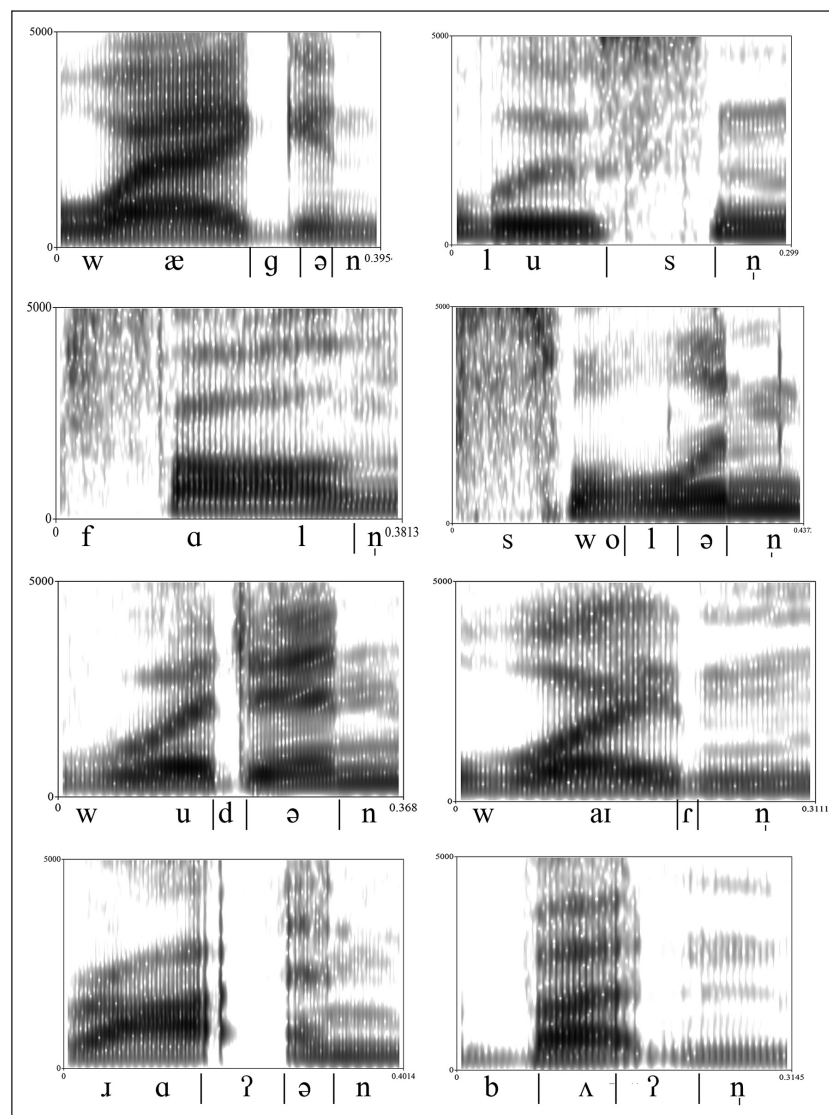

Figure 1: Example spectrograms of target words with [ən]/[n̩]. Each row, from left to right: wagon ([gən]), loosen ([sn̩]), fallen ([ln̩]), swollen ([lən]), wooden ([dən]), widen ([ɾn̩]), rotten ([ʔən]), button ([ʔn̩]). Vertical lines between the IPA symbols indicate the segmentation of the adjacent sounds for the [Cən] or [Cn̩] portion of the word. No segmentation between [ɑ] and [l] in 'fallen' is given, because it is not entirely clear where the boundary should be.
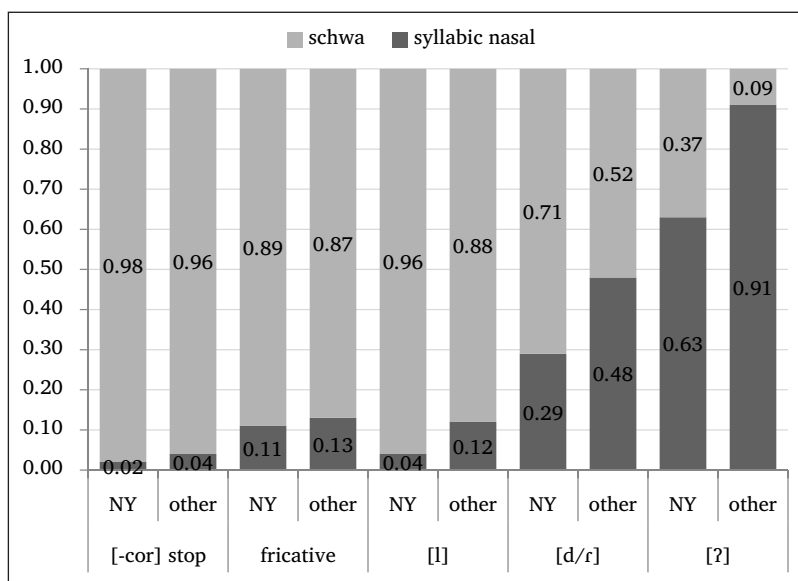
Figure 2: Proportion of schwas versus syllabic nasals by preceding consonant for laboratory study, divided by New York speakers versus speakers from other U.S. regions.

Table 2: Logistic regression results for effect of preceding consonant on the probability of schwa presence.

	Estimate	*z* value	Pr(>|z|)
(Intercept)	2.637	5.007	<0.001*
[-cor] stop	2.895	4.885	<0.001*
[l]	2.111	3.814	<0.001*
[d/ɾ]	−1.491	−3.219	0.001*
[ʔ]	−4.738	−8.573	<0.001*
non-NY	−0.828	−1.467	0.142
[-cor] stop:non-NY	0.547	1.166	0.244
[l]:non-NY	−0.207	−0.548	0.584
[d/ɾ]:non-NY	−0.020	−0.075	0.940
[ʔ]:non-NY	−1.008	−2.404	0.016*

Table 3: Tukey comparisons between preceding consonant contexts. With a Bonferroni correction (.05/10), significance is reached at .005.

	Estimate	*SE*	*z* value	Pr(>|z|)
fricative – [d/ɾ]	2.577	0.734	3.51	0.004*
stop – [d/ɾ]	4.386	0.860	5.099	<0.001*
[ʔ] – [d/ɾ]	−3.247	0.733	−4.432	<0.001*
[ʔ] – fricative	−5.825	0.823	−7.082	<0.001*
[ʔ] – [l]	−6.849	0.906	−7.562	<0.001*
[ʔ] – stop	−7.633	0.947	−8.063	<0.001*
[l] – [d/ɾ]	3.602	0.816	4.415	<0.001*
[l] – fricative	1.025	0.812	1.262	0.712
[l] – stop	−0.784	0.901	−0.87	0.906
fricative – stop	−1.808	0.857	−2.11	0.213

A closer examination of both [d/ɾ] and [ʔ] reveals an unexpected pattern: It is not the case that all speakers produce a significant portion of the [d/ɾ] tokens and most [ʔ] tokens with a syllabic nasal. Rather, visual inspection of individual speakers demonstrates that except for a small number of people, the majority of speakers produce either <25% or >75% syllabic variants. This is shown in **Figure 3**. The speakers from New York are marked in the axis labels; while the non-New York speakers are concentrated in the >75% syllabic nasal category for [ʔ], the New Yorkers are spread throughout the range. For [d/ɾ], the speakers are more evenly divided. For both [ʔ] and [d/ɾ], New Yorkers are the majority of the speakers in the <25% syllabic nasal category.

When it comes to the pronunciation of the coronal allophones, a closer look at the relationship between [d/ɾ] production and schwa or syllabic variants shows that there is no relationship between when speakers produce [ɾ] or [d] and when they produce a schwa or a syllabic nasal. On the other hand, for the [ʔ] words, speakers are more likely to produce a schwa when it is accompanied by glottalization than when [ʔ] is produced as a glottal stop with a period of complete glottal closure. This is shown in **Table 4**.
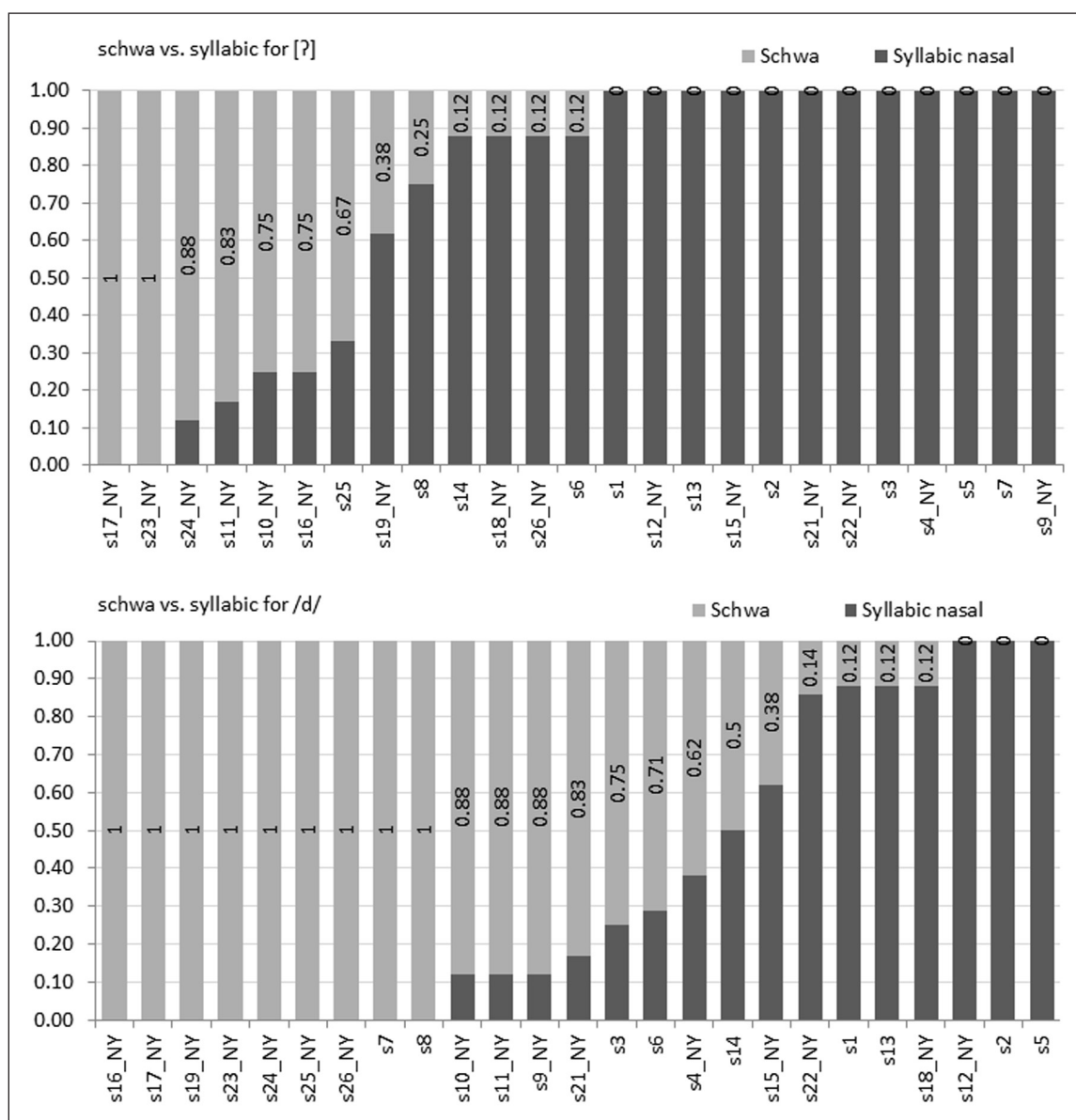

Figure 3: Proportion of schwa or syllabic nasal for preceding [ʔ] (top) and [d/ɾ] (bottom). Speakers from NY are marked.

Table 4: Relationship between implementation of [d/ɾ] and glottal stop/glottalization and realization as [ən] or [n̩].

		syllabic nasal	schwa
ex. *wooden, sudden*	[d] (37% of words)	0.41	0.59
	[ɾ] (63% of words)	0.35	0.65
ex. *button, eaten*	glottalization (47% of words)	0.64	0.36
	full closure (53% of words)	0.87	0.13

Using only the subset of the data including the [ʔ] and [d/ɾ] words since this is where the main variability occurred, we examined whether there is any relationship between lexical frequency and whether the words were produced as [ən] or [n̩]. A logistic mixed effects regression with preceding consonant ([ʔ] or [d/ɾ]), frequency calculated using the SUBTLEX$_{US}$ corpus (Brysbaert & New, 2009), and their interaction was carried out, with random intercepts for words and speakers. Results show that as expected, there is a main effect of preceding consonant (β = –4.84, z = –5.03, p < 0.001), but no main effect of frequency (β = –0.055, z = –1.70, p = 0.09) or interaction between preceding consonant and frequency (β = 0.051, z = 1.17, p = 0.24). While frequency is not significant, the negative coefficient could indicate a trend toward a greater proportion of schwas present when lexical frequency is lower if a bigger corpus were used. If the schwa variant is considered to be more hyperarticulated or less reduced, this trend is consistent with studies which have shown that more frequent words are often the more articulatorily reduced variants (e.g., Aylett & Turk, 2006; Gahl, 2008; Pluymaekers, Ernestus, & Baayen, 2005).

Lastly, we examine the duration of the nasal depending on whether it is syllabic or preceded by schwa. Results for this analysis are shown in **Tables 5** and **6**.

Table 5: Linear regression results for nasal duration in [ən] or [n̩].

	Estimate	SE	t value	Pr(>|t|)
(Intercept)	0.071	0.010	6.832	<0.001*
schwa present	−0.004	0.009	−0.435	0.663
[-cor] stop	0.004	0.013	0.296	0.768
[l]	−0.004	0.013	−0.296	0.768
[d/ɾ]	0.050	0.011	4.398	<0.001*
[ʔ]	0.031	0.011	2.760	0.006*
schwa:[-cor] stop	−0.001	0.011	−0.080	0.937
schwa: [l]	0.001	0.011	0.080	0.937
schwa:[d/ɾ]	−0.030	0.010	−3.114	0.002*
schwa:[ʔ]	−0.022	0.010	−2.225	0.026*

Table 6: Duration of the nasal in syllabic nasal productions and following schwa in [ən] productions. Standard deviations are in parentheses.

Preceding consonant	Syllabic nasal	SD	With schwa	SD
[-cor] stop	63	(23.5)	67	(24.7)
Fricative	78	(21.3)	71	(26.1)
[l]	56	(17.7)	65	(22.6)
[d/ɾ]	121	(35.9)	87	(25.4)
[ʔ]	101	(36.4)	79	(30.6)

The main effects for preceding consonant indicate that there is some variability in nasal duration across all of these contexts compared to the grand mean. The primary results of interest are the interactions between schwa presence and absence and preceding context, which shows that the syllabic nasal is significantly longer than the nasal after schwa for [d/ɾ] and [ʔ]. For other preceding sounds, there are no significant differences.

2.1.5. Summary

The results from the production study in the lab demonstrate that syllabic nasal consonants are only consistently produced when the word contains [ʔ] as a preceding consonant. Rates of [n̩] after [d/ɾ] are also higher than preceding [-coronal] stop, fricative, or [l], but not as high as after [ʔ]. These patterns hold for both New Yorkers and the speakers from other regions, though New Yorkers overall produce more schwas, substantially so after [ʔ]. These results are consistent with the hypothesis suggested by previous observations that a syllabic coronal nasal should be more likely after a coronal sound, though the current acoustic results do not resolve whether [ʔ] that contains a period of closure is actually being produced as a glottally reinforced [tʔ], or as a true glottal stop (with no concomitant tongue tip raising). These results also show that the variability seen in the overall results in **Figure 2** primarily reflect across-speaker variability, not within-speaker variability, as illustrated in **Figure 3**. Likewise, the minimal frequency effect on how [ən]/[n̩] words are produced is consistent with the across-speaker variability shown in **Figure 2** being the most important factor in explaining any gradience in the aggregate data, though there may be a tendency toward greater rates of syllabic nasal as frequency increases. Finally, syllabic nasals are significantly longer than the nasal in [ən] for [d/ɾ] and [ʔ]. Since the latter two are the main environments where there are enough syllabic nasals produced for the effect to be robust, further discussion of the effect of nasal length will focus on these preceding consonants.

In the following section, we carry out a similar analysis of the rates of [ən]/[n̩] to investigate whether the patterns in the lab study, with its emphasis on the New York region, also extend to the Pacific Northwest and Northern Cities.

2.2. University of Washington/Northwestern University (UW/NU) Corpus Data

2.2.1 Participants

The UW/NU Corpus consists of 33 talkers from the Pacific Northwest (11M, 9F) and Northern Cities (7M, 6F) dialect regions reading the IEEE Harvard sentences. Of these, 20 talkers were used (8 NC, 12 PN), since these ones produced all of the sentences containing all of the words in **Table 6**.

2.2.2 Stimuli

From the transcripts of the sentences used in this corpus, we identified 24 unique words (nine of which appeared more than once, in two or three different sentences) that contained the same type of potential word-final syllabic nasals that were used in the laboratory study. These are shown in **Table 7**, and the sentences containing these words are in the Appendix. In addition to the preceding consonant categories from the laboratory study, three words from this corpus had a preceding nasal. In total, 1153 words were analyzed. Note that since we used all possible words with final potential syllabic nasals, there are some instances here where the target word is the first or last in the sentence. Given the overall low rates of syllabic nasal production in the lab study, we decided to maximize the number of words rather than exclude these initial and final words. An examination of the results indicates that words in these positions had the same schwa/syllabic nasal profiles as the other words in their preceding consonant category.

Table 7: Words found in the UW/NU corpus. If more than one token of a word appeared in the sentences, it is noted in parentheses.

[-cor] stops	fricatives	[l]	[ʔ]	[d/ɾ]	nasal
open (×3)	often (×3)	fallen	button (×2)	wooden (×2)	woman
chicken (×2)	seven (×3)	sullen	kitten	hidden	women
wagon	even		eaten		salmon
broken	woven		frighten		
	fasten (×2)				
	person (×2)				
	cushion				
	frozen				
	season				

2.2.3 Results

Results were analyzed using a mixed effects logistic regression. The dependent variable was syllabic nasal or schwa responses, with preceding consonant category ([-cor] stop, fricative, nasal, [l], [d/ɾ], and [ʔ]) and speaker dialect as fixed effects and speaker and word as random intercepts. The interaction between preceding consonant category and dialect was also included in the model. Both factors were sum coded.

The proportions of syllabic nasals versus schwas for UW/NU data are shown in **Figure 4** and statistical results are given in **Tables 8** and **9**. Results show that schwas are significantly less likely for [ʔ] and [d/ɾ], and more likely for nasals and [-coronal] stops as compared to the grand mean. Tukey tests confirm that there are significant differences between [ʔ] and all other preceding consonant contexts, except [d/ɾ]. Syllabic nasals are also more likely for preceding [d/ɾ] than for fricatives, nasals, and stops. No other comparisons are significant. There is no main effect of dialect region, but the interactions show that there
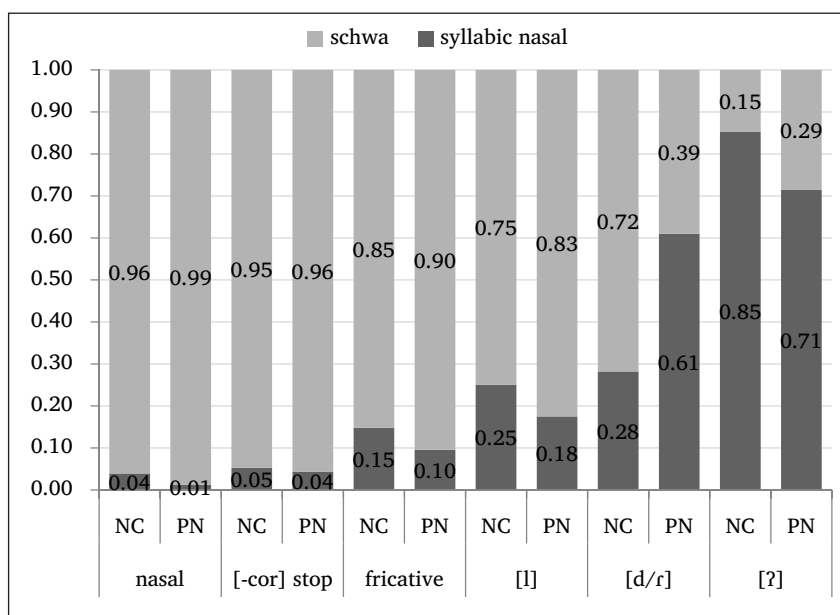

Figure 4: Proportion of schwas versus syllabic nasals by preceding consonant for the UW/NU corpus. NC: Northern Cities, PN: Pacific Northwest.

Table 8: Logistic regression results for effect of preceding consonant and dialect region on the probability of schwa presence.

	Estimate	z value	Pr(>|z|)
(Intercept)	2.363	7.471	<0.001*
[-cor] stop	1.436	2.871	0.004*
[l]	−0.441	−0.631	0.528
[d/ɾ]	−2.286	−3.603	<0.001*
[ʔ]	−3.948	−8.150	<0.001*
nasal	2.417	3.187	0.001*
region_PN	0.188	0.992	0.321
[-cor]stop:region_PN	−0.115	−0.483	0.629
[l]:region_PN	0.063	0.223	0.824
[d/ɾ]:region_PN	−0.849	−4.094	<0.001*
[ʔ]:region_PN	0.188	0.977	0.328
nasal:region_PN	0.256	0.578	0.564

Table 9: Tukey comparisons between preceding consonant contexts. With a Bonferroni correction (.05/15), significance is reached at .0033.

	Estimate	SE	z value	Pr(>|z|)
fricative − [d/ɾ]	2.8173	0.7419	3.797	0.0019*
[l] − [d/ɾ]	1.8449	0.957	1.928	0.3731
nasal − [d/ɾ]	4.7028	1.0571	4.449	<0.001*
stop − [d/ɾ]	3.7218	0.8667	4.294	<0.001*
[ʔ] − [d/ɾ]	−1.6624	0.7991	−2.08	0.2874
fricative − [ʔ]	4.4796	0.6143	7.293	<0.001*
[l] − [ʔ]	3.5073	0.8596	4.08	<0.001*
nasal − [ʔ]	6.3651	0.9734	6.539	<0.001*
stop − [ʔ]	5.3841	0.7614	7.072	<0.001*
[l] − fricative	−0.9723	0.7835	−1.241	0.8088
nasal − fricative	1.8855	0.8914	2.115	0.2693
nasal − [l]	2.8579	1.0842	2.636	0.0848
stop − fricative	0.9045	0.6599	1.371	0.735
stop − [l]	1.8768	0.9021	2.08	0.287
stop − nasal	−0.393	1.134	−0.346	0.999

is a significant effect for [d/ɾ] only: Pacific Northwest speakers are more likely to have syllabic nasals in this environment than Northern Cities speakers are. As a reminder, the overall proportion of syllabic nasals for [d/ɾ] for the New York speakers is 29%, and 48% for speakers from other regions (see **Figure 2**), which are both lower than the Pacific Northwest speakers in this study.

Since these results are based on read speech, the next section reports on data from a smaller corpus study of spontaneous speech as a preliminary attempt to investigate whether potential syllabic nasals are implemented similarly in read and spontaneous speech.

2.3. Spontaneous speech data

The stimuli for the spontaneous speech analysis were taken from the Fisher CTS corpus (Cieri et al., 2004), which was created in 2003. This corpus contains conversational telephone speech between two participants previously unknown to one another who are given a prompt to talk about. The corpus was developed to assist with the development of automatic speech recognition. Because the goal of this portion of the study is to get a general overview of spontaneous speech at large and whether it mirrors the read speech patterns, the speakers are not controlled for age, sex, or regional background. In the corpus overall, 38% of speakers are 16–29, 45% are 30–49, and 17% are over 50. Female speakers comprise 53% of the talkers. Speakers in the corpus are mostly evenly split over four broad geographic regions: North, Midland, South, West. Words were taken from speakers across all of the geographic regions.

Of the 40 words that were used in the laboratory study, those words which were produced at least 8 times in the Fisher CTS corpus were chosen. If there were fewer than 10 tokens in the corpus for any word, all of the tokens were used. If there were >10 tokens, a random sampling of 10 tokens were extracted from the corpus. This resulted in 230 utterances representing 25 of the 50 words from the lab study (Ns = [-cor] stop: 36, fricative: 44, [l]: 48, [d/ɾ]: 37, [ʔ]: 65). The same criteria for identifying schwa or syllabic nasal variants used in the lab study were also applied to the Fisher CTS data.

Results for this data are very similar to those for the lab study, except that the proportion of syllabic nasals produced after [d/ɾ] now matches the proportion for a preceding [ʔ]. This is shown in **Figure 5**. The same logistic regression and Tukey tests described in Section 2.1.4 indicate that there are significant differences for [d/ɾ] and [ʔ] with all other preceding consonants (both $p < 0.001$), but no significant difference between [d/ɾ] and [ʔ] or among any of the other preceding consonants.

The data from spontaneous speech suggest that the same pattern is largely found as for read speech, though the proportion of syllabic nasals for the [d/ɾ] words is now in line with that for [ʔ], which is very high. It may be that if there is more reduction or overlap in spontaneous conversational speech than in read speech, the [n̩] is even more likely to surface, though notably, only where it is already permitted; rates of schwa do not go down in the rest of the environments.
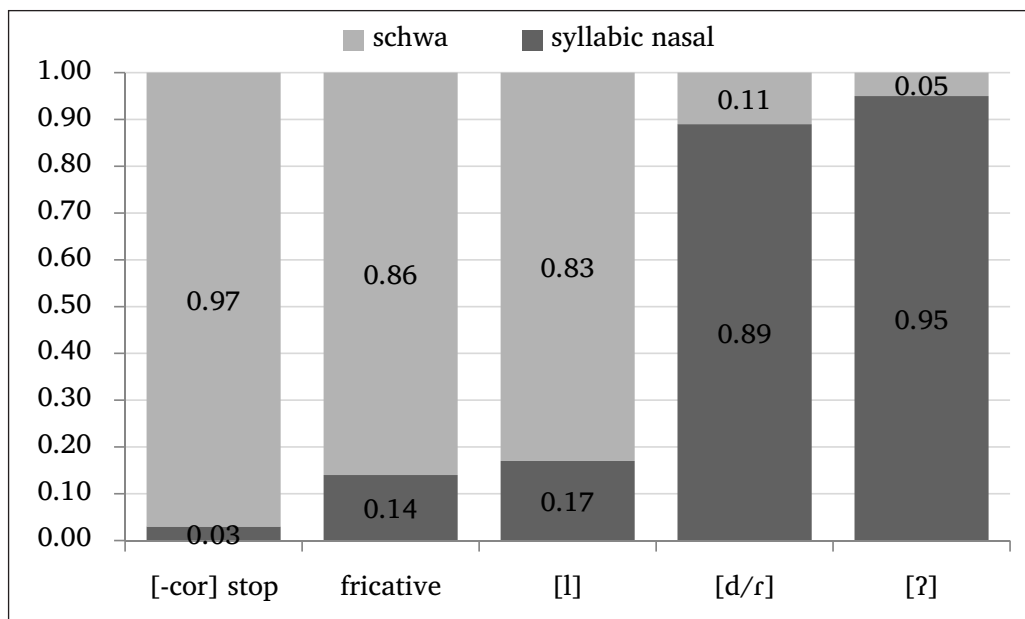

Figure 5: Proportion of schwas versus syllabic nasals by preceding consonant for Fisher CTS data.

3. General Discussion

The outcomes for the laboratory study, the UW/NU data, and the Fisher CTS spontaneous speech corpus all converge on very similar results. True syllabic nasals are most common after the [ʔ] allophone of /t/ that is almost obligatorily produced in many American English varieties and in these studies in this position.[2] In the read speech data, the rates of syllabic nasals after [ʔ] ranged from 63% to 85% across the groups examined in this study. A breakdown of the read data in the laboratory study, which was the best controlled for number of items in each preceding consonant category, showed that in fact these proportions reflect across-speaker variability, not within speaker. That is, some speakers produced almost all schwas after [ʔ], but most of the rest produced almost all syllabic nasals.[3] In spontaneous speech, the rates of syllabic nasal after [ʔ] are even higher, at 95%.

For most of the remaining preceding sounds in read speech, syllabic nasals are much less common, reaching 25% after laterals for the Northern Cities speakers, but not even reaching 20% for the other categories, for any other speakers. The exception to this is [d/ɾ], where speakers show intermediate rates of syllabic nasals (28–61%, depending on speaker group), but even this aggregate is misleading, since the speaker breakdown for the laboratory data shows that like [ʔ] words, speakers generally either mostly produce either syllabic nasals or schwas for [d/ɾ]. In this environment, more speakers produce [ən] than [n̩].

These results make it clear that syllabic nasals in word-final position in American English are limited in their distribution. They do not occur in most of the preceding consonant environments that have been provided as examples in many previous studies (see Section 1). On the other hand, the findings are consistent with the few papers that explicitly suggest that syllabic nasals should be most likely after coronal stops (Roach, 2009; Wells, 1995), and with the one empirical study of British English that showed this for a small number of speakers (Toft, 2002). As some have described (Carley et al., 2017; Roach, 2009; Zue & Laferriere, 1979), at first glance, a syllabic nasal following a coronal stop is articulatorily sensible, since the tongue can simply stay at the alveolar ridge during the transition from the oral to the nasal stop; this is what some authors have termed 'nasal release.' However, an account of where American English speakers produce syllabic nasals does not end here, because it is intertwined with the question of why [ʔ] might have arisen in American English in precisely this environment.

3.1. Variation and implementation in syllabic nasals

The general assumption that the default variant of [ən]/[n̩] is [ən], and that it can productively combine with many different sounds, is credible for at least two interrelated reasons. First, except in the [ʔ] and sometimes the [d/ɾ] environments, this study demonstrates that potential word-final syllabic nasals are actually produced as [ən]. Second, evidence from morphemic composition shows that [ən] is a productive word-final morpheme that is clearly realized with a schwa in many environments (e.g, the deadjectival morpheme, *ripe → ripen, damp → dampen,* past participles like *broke → broken, choose → chosen*, the demonymic morpheme, *Mexico → Mexican, Chile → Chilean*). Given that syllabic nasals are not realized in the large majority of environments in which they could occur, it is parsimonious to hypothesize that the default variant of these relatively

---

[2] We are aware that there are dialects of North American English that have flap as a variant of /t/ in the potential syllabic nasal words (e.g., mitten [mɪɾən]), though these have not been described in a paper to our knowledge. No speakers in any part of this study produced this variant, so the discussion in this paper applies to those dialects that have glottalization before syllabic nasals.

[3] Although the sample size for [ʔ] in the UW/NU corpus is much smaller, the same pattern of across-speaker variability holds. Only one Northern Cities speaker produces 40% of utterances with a schwa, with the remaining seven speakers at 20% or below. For Pacific Northwest speakers, three produced between 80–100% of [ʔ] words with a schwa, three speakers were between 40–60%, and the remaining six were at 20% or below.

productive morphemes is [ən], and that [ʔ] and [d/ɾ] are the special environments where something happens.

As a preliminary step, we argue that the original rise of the syllabic nasal in this environment is likely a result of a type of articulatory merging, along the lines of the descriptions of nasal release of a coronal stop followed by a coronal nasal (Carley et al., 2017; Roach, 2009; Zue & Laferriere, 1979). When /ən/ follows /t/, if the speaker leaves the tongue tip raised after producing a coronal stop in anticipation of the coronal nasal, then even if a reduced vowel gesture is produced by the tongue body, its realization would be masked if it is overlapped by the raised tongue tip. Such a configuration could occur for either /t/ or /d/, which are the two environments where syllabic nasal most often occurs. We return to the difference in rates between these two stops below. Moreover, a configuration in which the tongue tip stays raised for both a coronal stop and for /n/ is potentially compatible with a longer duration of the syllabic nasal, if speakers anticipatorily lower the velum as they would if the schwa were realized preceding a nasal (Bell-Berti, 1993; Cohn, 1993; Solé, 1995). That is, a portion of the tongue tip closure that should correspond to an oral stop might be realized as a nasal if the velum is anticipatorily lowered before the /t/ or /d/ itself is completed, which would acoustically turn the oral stop into a nasal one. Note that the duration results are consistent with Stevens and Keyser's (2010) speculation that syllabic nasals arise from the co-production of a tongue raising gesture for a coronal consonant and a tongue body gesture for a schwa. They also claim that the formant structure of syllabic nasals might be likewise affected, which we leave for future research.

In contrast, preceding sounds that do not have a coronal closure do not mask the schwa, since the articulatory advantage that results from not having to disrupt the tongue tip closure disappears if the closures preceding the nasal are produced with different articulators. It appears that even the coronal fricative articulation as in /s/ and /z/ is either not enough of a tongue tip constriction to justify leaving the tongue tip in place, or if speakers have differences in constriction locations for coronal fricatives and stops, as has been found for English speakers (Dart, 1998), then again, the advantage of raising the tongue tip to the same position that is required for /n/ is precluded.

Since individual speakers tend to be more or less categorical in implementing either schwa or syllabic nasal for [ʔ] and [d/ɾ], such a pattern suggests that the articulatory scenario described above is the phonetic precursor to what gave rise to the syllabic nasal variant following coronal stops, but not before other sounds. That is, today, [n̩] may be an allomorph of [ən] after coronal stops for those speakers who produce it consistently. It is not likely that [n̩]-speakers are intending to produce [ən] but due to common gestural overlap and reduction patterns in connected speech, end up with [n̩] every time; if this were the case, we should expect to see more variability within speakers as to whether they produce [ən] versus [n̩]. Instead, the use of the [n̩] variant may be dialectally conditioned, with New York and Pacific Northwest speakers having slightly lower rates than Northern Cities speakers and the other, geographically varied speakers in the laboratory study (though the data in this paper cannot speak to what might be conditioning these rates of allomorph choice). Note that we do not intend to rule out varying rates of gestural overlap entirely; we are not arguing that it is impossible for the [ə] to be obscured as a result of gradient overlap. To the extent that we do see any variability within speakers, gestural overlap could be playing a role. Nevertheless, if [n̩] is the variant that follows coronal stops, it may be easier to explain the [ʔ] allophone in American English. This is taken up in the next section.

3.2. Distinguishing between /t/ and /d/: [ʔ] before [n̩] in American English

We turn here to the realization of /t/ as [ʔ] before [ən]/[n̩] (Harris, 1994; Kahn, 1980). In one scenario, if [ən] is a morpheme that attaches to a lexical base that ends in /t/ (e.g., *eaten* /it + ən/), it could be expected that the form should surface as [iɾən]. Indeed, the ˈV

_ V stress pattern of such a word is precisely an environment where flapping in American English typically occurs (e.g., Borowsky, 1986; Kahn, 1980; Patterson & Connine, 2001; Turk, 1992; Warner & Tucker, 2011), and with other unstressed morphemes, such as /-ɪŋ/ or /-ɚ/, flapping does occur (*eating* [iɾɪŋ], *eater* [iɾɚ]). Why then, in the case of /ən/, does it not usually result in flapping when it is attached to a stem that ends in /t/? Where does the glottal variant of /t/ come from in this environment?

A possible answer to these questions may be found in accounts of glottal reinforcement of /t/ in coda position more generally that attribute glottalization to acoustic enhancement (Keyser & Stevens, 2006; Pierrehumbert, 1995; Seyfarth & Garellek, 2015; Stevens & Keyser, 2010; but see counterarguments in Huffman, 2005; Seyfarth & Garellek, 2020). In this line of research, coda /t/ glottalization is argued to be an acoustic modification that is implemented in order to enhance a phonological contrast that is otherwise perceptually endangered in a particular environment. For example, Keyser and Stevens (2006) hypothesize that in English, there is impetus to cut off vocal fold vibration in voiceless coda stops that are preceded by a vowel so that too much carryover phonation is not produced during the stop closure. They argue that in labial and velar stops, the tongue surface posterior to the constriction can be stiffened to help inhibit the expansion of the vocal tract volume (Svirsky et al., 1997), but this is more difficult in coronal stops which must have a more flexible tongue body in order to make a tongue tip closure. Instead, in this environment, another way to suppress vocal fold vibration is to adduct the vocal folds for the purpose of either glottalization or a glottal stop if there is full adduction. Keyser and Stevens propose that this is an example of featural enhancement, and that it explains why glottalization is more common for /t/ than for other voiceless stops, and why it occurs in coda position.

Since American English speakers may already implement this method to distinguish /t/ from /d/ in codas, this articulation could have been appropriated for word-medial position to also enhance the distinction between coda /t/ and /d/ before [ən]/[n̩]. If words with /t/ are produced with a glottally reinforced [tʔ], then it would be clear to listeners that the speaker is intending to produce /t/ even though at the same time, they are not lowering their tongue tip between the closure for the stop and the closure for the subsequent nasal. Since a neighboring nasal consonant may be a particularly good aerodynamic environment for bleeding voicing into an adjacent stop (Davidson, 2016), glottalization would also be a way to prevent this from happening and signal to the listener that the speaker intended a /t/ instead of a /d/. Moreover, glottal reinforcement for /t/ is particularly prevalent before sonorant consonants in American English (either word-finally at word boundaries, or in words where it is unambiguously a medial coda, as in *oatmeal* or *catnip*) (Huffman, 2005; Pierrehumbert, 1994; Seyfarth & Garellek, 2015, 2020) so the potential syllabic nasal environment is a natural extension of where [tʔ] would be expected.

To make an acoustic contrast with [tʔ], underlying /d/ can either be produced as a short ballistic [ɾ] which requires moving away from the alveolar ridge, or as a [d] with a lesser degree of overlap between the tongue tip gestures of [d] and [n], both of which are different from a /t/ that is realized with glottalization. Either way, [ə] is more likely to be realized for /d/ than for /t/, as reported in **Table 3**. The presence of the vowel is potentially another cue that could help distinguish between /t/ and /d/, which may explain the differences between rates of [n̩] after these coronal stops. However, at the same time, there may be pressure for the [n̩] allophone to spread from /t/ to the other coronal stop, since rates of [n̩] are higher for [d/ɾ] than for any of the other remaining preceding consonants in these studies. While the spontaneous speech data from the Fisher corpus has a small number of tokens, in this data [ʔ] and [d/ɾ] have similar high rates of [n̩] following these consonants. Whether this is because articulatory phonetic processes like overlap and reduction increase in spontaneous speech or because these speakers simply have higher rates of [n̩] is unclear, but this question could be further pursued with a larger data set.

To briefly return to the comparison between environments for flapping (e.g., ea[ɾ]er 'eater,' wi[ɾ]er 'wider') and those for glottalization (e.g., ea[ʔ]en 'eaten,' but wi[ɾ]en 'widen') in American English, it is notable that acoustic contrast is maintained on the consonant for /t, d/ preceding the syllabic nasal, but both of those sounds neutralize to [ɾ] in other environments. Some research has indicated that where flapping occurs, speakers do produce slightly longer vowels before /d/ flaps than before /t/ flaps (Braver, 2014; Herd et al., 2010; Zue & Laferriere, 1979), and there are differences in F1 and F2 for the specific *writer/rider* pair (Kwong & Stevens, 1999), suggesting that the acoustic contrast is pushed to the vowel where flapping occurs. However, the same studies also show that listeners cannot productively use this length different to reliably distinguish /d/ flaps from /t/ flaps (Braver, 2014; Herd et al., 2010), so it remains an open question why American English maintains such a salient acoustic contrast between /t/ and /d/ in pre-nasal position, but apparently less so in the flapping environment. Anecdotally, some speakers of American English may have begun to produce [ʔn̩] as [ɾən] (e.g., *kitten* as [kɪɾən]), which could lead to the elimination of [n̩] as an environment where /t/ is produced as something other than a flap in intervocalic position. However, the extent of this change is currently unknown and will need to be revisited.

3.3. Acoustic and individual variability in the implementation of [ʔn̩]/[ʔən]

At this point, the phonetic account for syllabic nasals contains the following elements: Voiceless coronal stops before [ən]/[n̩] are realized with glottal reinforcement for acoustic enhancement, and there is greater prevalence of the [n̩] variant after these stops, which possibly originated from the oral tongue tip articulations for the /t/ and /n/ merging and masking the /ə/. This account, however, is only adequate for cases where speakers are raising their tongue tips for [tʔ] (which is probably usually realized with glottalization on the preceding vowel) but it is likely that some speakers are actually producing a full glottal stop that is not accompanied by any oral closure gesture at all. For the cases that were classified as full glottal stops in the laboratory study, it is not possible to know for certain whether a period of closure contains both an oral and a glottal closure, or only a glottal closure. At the same time, 47% of all of the [ʔ] utterances in the laboratory study were produced as a period of glottalization that had no evident closure (see **Table 3**). The distinction between a /t/ that is produced as only glottalization or with a period of closure is shown in **Figure 6**. Interestingly, speakers who produced a period of glottalization with no closure are significantly more likely to also produce a schwa (see **Table 3**), which is what might be expected when the tongue tip is not raised, since the incentive to keep the tongue tip raised from /t/ to /n/ is not present.
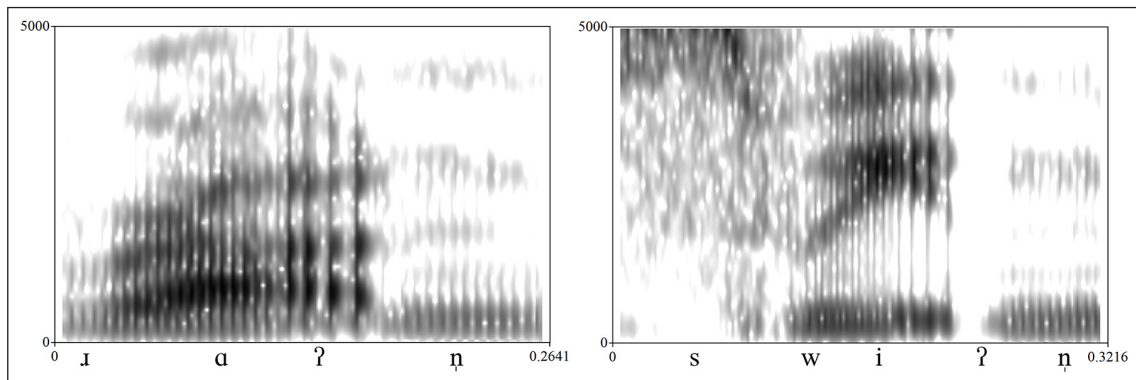

Figure 6: Examples (from the same speaker) of /t/ produced as a period of glottalization (*rotten*, left) and as a full glottal stop (*sweeten*, right) before [n̩].

Ultimately, if speakers variably implement a full glottal stop (with no simultaneous tongue tip closure), a period of glottalization, or glottally reinforced [tˀ], this may be indicative of speakers' knowledge that these realizations are not contrastive in English (e.g., Dilley, Shattuck-Hufnagel, & Ostendorf, 1996; Docherty & Foulkes, 1995, 2005; Pierrehumbert, 1995; Pierrehumbert & Talkin, 1992; Sumner & Samuel, 2005). If there are speakers who implement a syllabic nasal following either glottalization or a full glottal stop, or both, this is consistent with the point that the phonetic sketch delineated at the top of this section may have been the articulatory impetus for the rise of the glottal stop allophone of /t/ preceding a syllabic nasal allomorph in American English, which has become conventionalized and now may be realized with a broader set of articulatory options. The frequency with which speakers actually raise their tongue tips could be investigated with an imaging technique like electromagnetic midsagittal articulography (EMA), ultrasound, or real time MRI (rtMRI).

While [ʔn̩] is by far the most common implementation for words with /t/ as the underlying preceding consonant, the results from the laboratory study show that there are six New York speakers who produce [ʔən] more than 75% of the time. Three of these speakers produce a period of glottalization in 62% of the cases, two speakers produce only a period of glottalization, and one speaker produces only a period of closure (i.e., glottal stop or [tˀ]). If productions with a period of closure do reflect [tˀ] with a tongue tip raising component, then presumably these speakers' coordination pattern has the tongue tip releasing after the [tˀ] before the subsequent [n] on the tongue tip tier, which means that the [ə] would be acoustically realized. For the cases when speakers produce a period of glottalization, then this would be timed to end before the vowel is completed. This may be a dialectal variant, such as for the New York speakers in this study, and perhaps an ongoing sound change for some areas, given Eddington and Savage's (2012) finding that [ʔən] is more common among young female speakers in Utah than the non-Western U.S. speakers in their study.

4. Conclusion

Despite the relatively expansive assumptions about the distribution of the syllabic nasal in word-final position in English in the literature, empirical evidence from a laboratory study, a corpus of read sentences spanning multiple dialect regions, and samples of spontaneous speech indicate that syllabic nasals in American English only occur reliably after a glottal stop, a period of glottalization, or glottally reinforced [tˀ], and to a lesser extent, following /d/, which can be realized as either [d] or [ɾ]. The results also indicate that there is dialectal variation in the rates of [ən] versus [n̩], with New Yorkers being the most likely to retain [ən], especially for /ʔ/, as compared to Northern Cities and the Pacific Northwest. Data for individual speakers show that most speakers prefer either [ən] or [n̩], with some variation across speakers but little variation within speakers.

Given the widespread presence of [ən], and a feasible articulatory account for why the continued raised tongue tip from a preceding coronal stop or flap would acoustically mask the [ə], giving rise to a [n̩], it is parsimonious to assume that [ən] is the default variant of the [ən]/[n̩] morpheme (when it is morphemic). Also, it is likely that [n̩] is an allomorph that speakers select rather than regularly generate via articulatory processes like connected speech overlap or reduction every time. We then considered why [ʔ] is the surface realization in American English instead of either [t] or [ɾ], since the combination of a stem and suffix like /ɹɑt + ən/ for *rotten* is an environment in American English where flapping is expected. This was explained by extending to pre-syllabic nasal position an acoustic enhancement account that has been proposed for why /t/ is glottally reinforced in coda position. Glottalization, whether in conjunction with [t], or as a full glottal stop

or a period of creakiness, helps to distinguish /t/ from /d/ in the pre-nasal environment by implementing an articulatory variant associated with /t/ that already exists in other positions.

Additional File

The additional file for this article can be found as follows:

· **Appendix:** The sentences read by the participants in the laboratory study. DOI: https://doi.org/10.5334/labphon.224.s1

Acknowledgements

We would like to thank members of the NYU PEP lab and audiences at the University of Pennsylvania, Brown University, Scott Seyfarth and Marc Garellek, and the Linguistics Society of America meeting in January 2019 for their feedback on this work. We also thank Mark Liberman for his comments and for helping us access the Fisher CTS Corpus.

Competing Interests

The authors have no competing interests to declare.

References

Aylett, M., & Turk, A. (2006). Language redundancy predicts syllabic duration and the spectral characteristics of vocalic syllable nuclei. *The Journal of the Acoustical Society of America, 119,* 3048–3059. DOI: https://doi.org/10.1121/1.2188331

Bates, D., Maechler, M., Bolker, B., Walker, S., Christensen, R., Singmann, H., … Fox, J. (2018). lme4: Linear mixed-effects models using Eigen and S4. *R package version 1.* 1–19. Retrieved from http://CRAN.R-project.org/package=lme4

Bell-Berti, F. (1993). Understanding velic motor control: Studies of segmental context. In M. Huffman & R. Krakow (Eds.), *Phonetics and Phonology: Nasals, Nasalization and the Velum* (Vol. 5). New York: Academic Press. DOI: https://doi.org/10.1016/B978-0-12-360380-7.50007-7

Boersma, P., & Weenink, D. (2020). *Praat: Doing phonetics by computer [Computer program].* version 6.1.16.

Borowsky, T. (1986). *Topics in the Lexical Phonology of English.* (Doctoral dissertation). Amherst: University of Massachusetts.

Braver, A. (2014). Imperceptible incomplete neutralization: Production, non-identifiability, and non-discriminability in American English flapping. *Lingua, 152,* 24–44. DOI: https://doi.org/10.1016/j.lingua.2014.09.004

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

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

Carley, P., Mees, I., & Collins, B. (2017). *English Phonetics and Pronunciation Practice.* London: Taylor and Francis. DOI: https://doi.org/10.4324/9781315163949

Cieri, C., Graff, D., Kimball, O., Miller, D., & Walker, K. (2004). Fisher English Training Speech Part 1 Speech LDC2004S13. Philadelphia: Linguistic Data Consortium.

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

Cruttenden, A. (2008). *Gimson's Pronunciation of English* (7th ed.). London: Hodder Education.

Dart, S. (1998). Comparing French and English coronal consonant articulation. *Journal of Phonetics, 26*(1), 71–94. DOI: https://doi.org/10.1006/jpho.1997.0060

Davidson, L. (2016). Variability in the implementation of voicing in American English obstruents. *Journal of Phonetics, 54,* 35–50. DOI: https://doi.org/10.1016/j.wocn.2015.09.003

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

Docherty, G., & Foulkes, P. (1995). Acoustic profiling of glottal and glottalised variants of English stops. *Proceedings of the XIIIth International Congress of Phonetic Sciences,* 350–353.

Docherty, G., & Foulkes, P. (2005). Glottal variants of (t) in the Tyneside variety of English: An acoustic profiling study. In W. Hardcastle & J. Beck (Eds.), *A Figure of Speech – a Festschrift for John Laver* (pp. 173–199). London: Lawrence Erlbaum.

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

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

Fukaya, T., & Byrd, D. (2005). An articulatory examination of word-final flapping at phrase edges and interiors. *Journal of the International Phonetic Association, 35*(1), 45–58. DOI: https://doi.org/10.1017/S0025100305001891

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

Gussman, E. (1991). Schwa and syllabic sonorants in a non-linear phonology of English. *Acta Universitatis Wratislaviensis 1061: Anglica Wratislaviensia XVII,* 25–39.

Hall, T. A. (2006). English syllabification as the interaction of markedness constraints. *Studia Linguistica, 60*(1), 1–33. DOI: https://doi.org/10.1111/j.1467-9582.2006.00131.x

Hammond, M. (1999). *The Phonology of English.* Oxford: Oxford University Press.

Harris, J. (1994). *English Sound Structure.* Oxford: Blackwell.

Herd, W., Jongman, A., & Sereno, J. A. (2010). An acoustic and perceptual analysis of /t/ and /d/ flaps in American English. *Journal of Phonetics, 38,* 504–516. DOI: https://doi.org/10.1016/j.wocn.2010.06.003

Heselwood, B. (2007). Schwa and the phonotactics of RP English. *Transactions of the Philological Society, 105*(2), 148–187. DOI: https://doi.org/10.1111/j.1467-968X.2007.00186.x

Hothorn, T., Bretz, F., & Westfall, P. (2008). Simultaneous inference in general parametric models. *Biometrical Journal, 50*(3), 346–363. DOI: https://doi.org/10.1002/bimj.200810425

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

Kahn, D. (1980). *Syllable-based generalizations in English phonology.* New York: Garland.

Keyser, S. J., & Stevens, K. (2006). Enhancement and overlap in the speech chain. *Language, 82*(1), 33–63. DOI: https://doi.org/10.1353/lan.2006.0051

Kuznetsova, A., Brockhoff, P. B., & Christensen, R. (2013). lmerTest: Tests for random and fixed effects for linear mixed effect models (Version 2.0–33) [R package].

Kwong, K., & Stevens, K. (1999). On the voiced-voiceless distinction for writer/rider. *Speech Communication Group Working Papers, MIT Research Laboratory of Electronics*, *11*, 1–20.

Ladefoged, P., & Johnson, K. (2014). *A Course in Phonetics* (7th ed.). Stamford, CT: Cengage.

Mora Bonilla, J. (2003). The formation of syllabic consonants and their distribution in Southern British English. *Atlantis, 25*(2), 97–112.

Panfili, L. M., Haywood, J., McCloy, D. R., Souza, P. E., & Wright, R. A. (2017). *The UW/NU Corpus, Version 2.0*. Retrieved from: https://depts.washington.edu/phonlab/resources/uwnu/uwnu2/

Patterson, D., & Connine, C. M. (2001). Variant Frequency in Flap Production. *Phonetica, 58*(4), 254–275. DOI: https://doi.org/10.1159/000046178

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

Pierrehumbert, J. (1995). Prosodic effects on glottal allophones. In O. Fujimura & M. Hirano (Eds.), *Vocal fold physiology: Voice quality control* (pp. 39–60). San Diego: Singular Publishing Group.

Pierrehumbert, J., & Talkin, D. (1992). Lenition of /h/ and glottal stop. In G. Docherty & D. R. Ladd (Eds.), *Papers in Laboratory Phonology II: Gesture, Segment, Prosody* (pp. 90–116). Cambridge: Cambridge University Press. DOI: https://doi.org/10.1017/CBO9780511519918.005

Pluymaekers, M., Ernestus, M., & Baayen, R. H. (2005). Lexical frequency and acoustic reduction in spoken Dutch. *Journal of the Acoustical Society of America, 118*(4), 2561–2569. DOI: https://doi.org/10.1121/1.2011150

Polgárdi, K. (2014). Syncope, syllabic consonant formation, and the distribution of stressed vowels in English. *Journal of Linguistics*, *51*(2), 383–423. DOI: https://doi.org/10.1017/S0022226714000486

Recasens, D. (2012). A cross-language acoustic study of initial and final allophones of /l/. *Speech Communication*, *54*, 368–383. DOI: https://doi.org/10.1016/j.specom.2011.10.001

Roach, P. (2009). *English Phonetics and Phonology*. Cambridge: Cambridge University Press.

Roach, P., Sergeant, P., & Miller, D. (1992). Syllabic consonants at different speaking rates: A problem for automatic speech recognition. *Speech Communication*, *11*, 475–479. DOI: https://doi.org/10.1016/0167-6393(92)90054-B

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

Rubach, J. (1996). Shortening and ambisyllabicity in English. *Phonology, 13*(2), 197–237. DOI: https://doi.org/10.1017/S0952675700002104

Seyfarth, S., & Garellek, M. (2015). Coda glottalization in American English. In *Proceedings of the 18th International Congress of Phonetic Sciences*.

Seyfarth, S., & Garellek, M. (2020). Physical and phonological causes of coda /t/ glottalization in the mainstream American English of central Ohio. *Laboratory Phonology, 11*(1), 24. DOI: https://doi.org/10.5334/labphon.213

Shockey, L. (2003). *Sound Patterns of Spoken English*. Oxford: Blackwell. DOI: https://doi.org/10.1002/9780470758397

Simpson, A. (2005). 'From a Grammatical Angle': Congruence in Eileen Whitley's phonology of English. *York Papers in Linguistics, series 2*, 49–90.

Solé, M. J. (1995). Spatio-temporal patterns of velopharyngeal action in phonetic and phonological nasalization. *Language and Speech, 38*, 1–23. DOI: https://doi.org/10.1177/002383099503800101

Stevens, K., & Keyser, S. J. (2010). Quantal theory, enhancement and overlap. *Journal of Phonetics*, *38*, 10–19. DOI: https://doi.org/10.1016/j.wocn.2008.10.004

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

Svirsky, M. A., Stevens, K., Matthies, M., Manzella, J., Perkell, J., & Wilhelms-Tricarico, R. (1997). Tongue surface displacement during bilabial stops. *The Journal of the Acoustical Society of America, 102*(1), 562–571. DOI: https://doi.org/10.1121/1.419729

Szigetvári, P. (2002). Syncope in English. *The Even Yearbook*, *5*, 139–149.

Toft, Z. (2002). The phonetics and phonology of some syllabic consonants in Southern British English. *ZAS Papers in Linguistics, 28*, 111–144. DOI: https://doi.org/10.21248/zaspil.28.2002.162

Trager, G. L., & Bloch, B. (1941). The syllabic phonemes of English. *Language, 17*(3), 223–246. DOI: https://doi.org/10.2307/409203

Turk, A. (1992). The American English flapping rule and the effect of stress on stop consonant durations. *Working Papers of the Cornell Phonetics Laboratory, 7*, 103–133.

Warner, N., Fountain, A., & Tucker, B. (2009). Cues to perception of reduced flaps. *Journal of the Acoustical Society of America, 125*, 3317–3327. DOI: https://doi.org/10.1121/1.3097773

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

Wells, J. C. (1995). New syllabic consonants in English. In J. W. Lewis (Ed.), *Studies in General and English Phonetics* (pp. 401–412). London: Routledge.

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