York University LAL Major Research Paper (MRP)
Student No.: 205552583
Date: Monday, August 8th, 2022
Supervisor: Emily Elfner

The effects of prosody and production planning in external t-sandhi

Cydklaire Borja

Acknowledgements

It is entirely thanks to Emily Elfner, who provided MRP supervision in equal parts of theoretical, technical, and emotional support, that this project and paper were possible. Her guidance truly gave me the confidence to explore my ideas. I am also extremely grateful to Gabriela Alboiu, whose knowledge and encouragement have been vital to my decades long interest in linguistics.

I am indebted to the family and friends who either made up the body of my participants, or who were simply (but integrally) my supporters: thank you for allowing me to disappear for a year, resurfacing temporarily only to ask you to participate in a study, or to ramble about tiny puffs of air and deleted sounds.

At last, thank you to my partner Simon Baker, for allowing me to rely on his innumerable wonderful qualities and abilities this year, and my mother, Maria Balingit, whose courage and good humour helped me power through the tough spots.

1. Introduction

The change in pronunciation of word-final /t/ in English is known as external coronal stop sandhi and is understood to be motivated by several phonological conditions. Though much of the research in this area (e.g., Tanner et al., 2017, Kilbourne-Ceron et al., 2016, Coetzee and Kawahara, 2013) has generally focused on a single t-sandhi alternation within a corpus (e.g., from /t/ to /ɾ/), the results of these studies make it clear that it is always a combination of factors that play a role in each alternation. In this paper, I propose that examining the distribution of various types of t-sandhi clause-internally would provide a more complete look of where speakers preferred to use different forms of /t/ in their natural speech: tightly controlling an area of analysis would make this otherwise wide scope more manageable. This project uses two production experiments to examine the interaction between prosody and other conditioning factors on t-sandhi: established segmental contexts and findings from Production Planning studies were used to control the environment surrounding word-final /t/ and record the distribution of its pronunciation.

The results of these studies show that the more marked forms of t-sandhi (in terms of production) favor prosodically defined environments not explored in previous studies: word-external, clause-internal flapping exclusively preferred environments where the following word did not carry lexical stress on the first syllable; released-t's were never produced in environments where the /t/ was preceded by a vowel and followed by a consonant, and also displayed changes in production rates that were sensitive to prosodic correlates. Production planning variables also showed some expected results which show areas of interest for further t-sandhi and production experiments focused on the production planning hypothesis (Tanner et al. 2017, Kilbourne-Ceron et al. 2016).

2. <u>Background</u>

Variations of /t/ include t-deletion, flapping, glottalization, glottal stops, or released t's (also referred to as tense/aspirated stops) and though many of the conditions that produce these alternations internally are well known, with the exception of more recent publications studies of external t-sandhi within a clausal boundary are less common.

2.1 Segmental Environment

Many t-sandhi studies have focused on coronal stop deletion (CSD) in particular. A pivotal study in this area, Guy (1980), was a model for many that followed: the phonological contexts that condition CSD were described and explored alongside a complex non-phonological sociolinguistic analysis (Guy, 1980). Many of the segmental contexts described therein (that are now widely understood to strongly affect CSD) fall into one of three categories: (1) segments preceding /t/: e.g., segments that share articulatory features with /t/, like voicelessness or coronality, and complex clusters (like /kst/ in the word mixed, /mɪkst/) will promote deletion; (2) segments following the /t/: CSD is promoted following consonants and inhibited following vowels; and (3) morpho-phonological conditioning: deletion is less likely when /t/ is an -ed suffix. In Guy's (1980) study and many others, t-sandhi's conditioning contexts are focused within the prosodic word and though the aim of this study is to examine external t-sandhi, the conditioning effects of /t/'s segmental environment remain applicable and integral to the experiment design.

Other segmental t-sandhi conditions, which were not considered explicitly but which provided the motivation to attempt at balancing the vowel quality, include findings that have shown that glottalization (sometimes combined with glottal stops in analyses) occurred more frequently when preceded by a front vowel (Eddington & Taylor, 2009), and the segmental conditioning of flapping, which is slightly different from the other conditions in that it is an articulatory requirement of flapping to occur intervocalically.

2.2 Prosody

Guy (1980) ultimately chose not to include prosodic effects in his study, but did briefly describe one effect of prosody on CSD: when /t/ occurred in a stressed syllable it was less likely to undergo deletion. Flapping is also mostly understood in the context of word-internal prosody: Kilbourne-Ceron et al. (2016:2) refer to the flapping distribution of /t/ as categorical within a word, unless the following vowel is unstressed. Studies on released-t are less common but are understood to be more commonly pronounced in the onset of a syllable, at least word-internally, i.e., as an aspirated /t^h/ (Hoard, 1971).

Studies that have looked at external t-sandhi have often been focused on non-phonological effects, like frequency (Coetzee and Kawahara, 2013, Bybee and Scheibman, 1999), and/or probability, but there are some studies that have focused on the effect of boundary strength, a prosodic correlate which can indicate the presence of a syntactic boundary. In their corpus study on CSD, Tanner et al. (2017) examined the effect of boundary strength[1] as an indication of the availability of production planning. They found that boundary strength (indicated by pause length in this study) negatively affected deletion rate (Tanner et al. 2017:19), and greatly modulated the segmental effect of the following word: when the pauses were long enough (362ms) there was almost no segmental CSD effect from the following context (Tanner et al. 2017:27).

Pauses are only one measure of boundary strength: articulatory strengthening (Fougeron and Keating, 1997, and Cho and Keating, 2001), voice onset time, vowel lengthening (Keating et al., 2003) and word duration (Lehiste 1973, Elfner, 2016) are some other measures which have been argued to correlate with the strength of prosodic boundaries. Kilbourne-Ceron et al. (2016) examined vowel lengthening as a measure of boundary strength and found that while increased vowel length did reduce the likelihood of flapping, it was syntactic clausal boundaries that had a significant effect on flapping. In this study, the boundary strength measure required combination with a syntactic boundary (in the negative direction) to correctly model flapping predictability (Kilbourne-Ceron et al., 2016: 7). The interaction between pauses and syntax is not unexpected or particularly inseparable from prosody: Elfner's (2016) study showed that boundary strength, indicated by the duration of a word before the boundary, increased before a phrasal subject, providing evidence for phonological phrasing mapping to syntactic phrasing. These findings highlight the importance of controlling both syntax and boundary strength.

Prosodic elements are interactive, and elements lower in the prosodic hierarchy may be affected by elements in the higher layers: higher prosodic layers carry syntactic structure or information structure, via sentence melodies or phrasal intonation (Féry, 2016; 178, 225-283). Lower prosodic elements, like lexical stress (the assignment of prominence on a syllable, at the level of the prosodic word), may be anchor points for these melodies, but are still subject to alteration by them: phonological elements and processes can also be affected this way: the dissimilation rendaku rule in Japanese, for example, is sensitive to compounding of prosodic words within the prosodic phrase (Féry, 2016: 53). Exploring the role of t-sandhi and prosody within the clause would mean looking at lower-level prosodic constituents such as the syllable. In fact, the syllable is an already identified source of internal t-sandhi: the role of /t/ is sometimes described as ambisyllabic: that is, existing in both the coda of the first syllable, and in the onset of the second (Féry, 2016; 41); this account, which can also be thought of as re-syllabification, describes the motivation for (and not just a description of) the flapping of /t/, which may prove more useful in predictions. Something like lexical stress, which involves prosody at the syllable level, or boundary strength, which involves the intersection between prosodic boundaries, are possible sources of t-sandhi conditioning.

[1] Though Tanner et al. (2017), and Kilbourne-Ceron et al. (2016) also look at probability (Tanner et al., 2017) and frequency effects (both), boundary strength was also an area of considerable focus: these were considered indicators of the strength of the production planning hypothesis.

2.3 Production Planning Hypothesis

The Production Planning Hypothesis (PPH) (Tanner et al., 2016, and Kilbourne et al., 2016), posits that when planning a sentence, the conditioning effects of sandhi can only be applied if the upcoming word is known at the time of planning. Conditioning in the form of syntactic boundaries as in Kilbourne-Ceron et al. (2016), or boundaries represented by pauses as in Tanner et al. (2016), affect a speaker's ability to apply deletion or flapping rules. These studies used a measure of word frequency or probability[2] to explain some of the variation found in t-sandhi: words that were more frequent were more prone to final /t/-deletion, but as the frequency strength increased, the effect of segmental context was reduced (Tanner et al., 2016: p.24). They attributed this to the PPH, where they propose that higher frequency words would be planned earlier, before the following phonological context is available. In Kilbourne-Ceron et al. (2016) a reduction in flapping before a syntactic boundary is attributed to a decrease in the availability of upcoming segmental information, but while the production experiment in the study varies the clausal boundary, it doesn't vary the availability of the segmental information in any temporal/visual- with the whole sentence available to speakers, it's less clear that a reduction in flapping isn't a result of planning for a boundary which makes flapping difficult. However, the prediction made by the PPH, that is: "a probabilistic decrease in the availability of upcoming segmental information under increased planning load" (Kilbourne-Ceron et al. 2016: 8) remains compelling theoretically, given the ability of PPH studies to predict deletion or flapping with some significance.

2.4 Hypothesis and study design

What makes corpus data attractive is the large sample size, but a drawback to this type of study is that it can often result in imbalanced data (Schmitz et al. 2021, p.29). However, the downside of experimental data (in the traditional form of sentence elicitations) is that it is likely to have some effect on a process that is as variable as final-t sandhi is, especially considering the known effects that formal read speech has on t-sandhi (Coetzee and Kawahara, 2013: 62). In Schmitz et. al.'s (2021) study on morphemic length variation, rigorously controlled conditions were used to bridge the gap between the imbalance of corpus studies and contradictory results yielded by experiments with less robust controls. This approach will be used here in two production experiments, which will examine conditioning effects on the production of all types of t-sandhi, as well as possible effects of the PPH.

Experiment 1 aims to closely control phonetic correlates of stress and boundary strength through several means: a) target words will be controlled for syllable complexity and length,

---

[2] Frequency was measured in Tanner et al. (2017) as a log-transformed corpus frequency (tokens per million) of all t/d-final words in the corpus used in the study and probability was measured as a bigram of probability fitted to the British National corpus (Tanner et al 2017: 17).

and will be placed in the same syntactic positions throughout all items, b) the word following the /t/ will be emphasized, both through explicit direction to the participant, as well as by using a narrow focus phrase, which ensures that the focused constituent (in this case, the word following the /t/) receives metrical prominence (Féry, 2016; 230-231). Control of the syntactic and prosodic boundary allows for narrower variation of the boundary size between the word-final t and the following word, which was anticipated to increase the visibility of the influence of boundary strength or stress on various forms of t-sandhi. Though stress may rely on acoustic correlates, it remains an abstract quality which simply represents prominence (Féry, 2016; 178), and as such may present itself in a few ways phonetically: e.g., an increase in pitch, duration, or intensity, which are all also measures of boundary strength.

A second production experiment, Experiment 2, will explore the effect of the PPH through visual restriction of the target words, using the same controls for phonology and prosody in the first experiment.

These two experiments will test and make the following predictions in the following areas:

1. T-sandhi distribution: four categories of t-sandhi: flaps, releases, glottalizations and unpronounced /t/ will be annotated, to examine t-sandhi proportions, as well as any unexpected segmental or prosodic effects.
2. Prosodic conditioning: as stronger prosodic boundaries have been previously found to inhibit flapping, lexical stress that occurs adjacent to the word-final /t/ will be perceived as a stronger boundary which will inhibit resyllabification, and flaps will be less likely. The opposite result is expected in released-t pronunciations. Acoustic correlates for increased boundary strength (increased duration, intensity and/or pitch deltas) will be present where prosodic conditioning involving boundary strength is observed.
3. Phonological conditioning: phonological contexts known to influence certain types of t-sandhi will be observed: flaps will still only be able to apply intervocalically; phonological conditions that might have an effect on the lesser studied external sandhi like released /t/ and glottalization will be observed.
4. Production Planning Hypothesis: reducing access to the verb containing the word final-/t/, or the following word (and its following phonological context), will reduce the speaker's ability to apply t-sandhi rules, leading to an increase in the speaker's default pronunciation of /t/; but there would be a larger effect expected in the following word condition, as predicted by the PPH.

3. <u>Methodology</u>

3.1 Participants and recording

Ten participants, two men and eight women between the ages of 20 and 45 with no background in Linguistics, were recruited from the author's community. All participants were native speakers of Canadian English, living in Toronto, Ontario. None reported speech or reading disabilities.

Three speakers fluently spoke a language other than Canadian English, but only one of those speakers used this language more than once a week. Speech data was recorded using Audacity in a home office. Stimuli were presented in pseudo-randomized order using Gorilla Experiment Builder. Half the participants started with Experiment 1 and the other half with Experiment 2.

3.2 Speech Materials

A total of 81 unique elicitation sentences across both experiments were used: each sentence was structurally similar and began with a subject (a name) and was followed by a verb phrase. The target words were one syllable verbs ending in /t/, and they were all placed in the same position in the verb phrase: at the beginning of a verb phrase, which was the same length in syllables, to reduce pitch and duration variability. To control prosody, the word following the verb received a narrow focus reading via a contextual sentence and was presented in bold and italicized font. Contextual sentences were also unique. For example, in Table 1 (sentence h.), the contextual sentence for this elicitation was '*Thomas' mom asked if his renovation cost a lot.'*

Half of the verbs ended in a vowel followed by a /t/ (.e.g, *fit*), the other half ended in a consonant, followed by a /t/ (e.g., built): half of the consonant-/t/ tokens ended in /lt/, and the other half in /nt/. The words that followed the word-final t varied between vowel initial and consonant initial words: consonants were limited to /n/ and /m/ to facilitate pitch measurements and were balanced between the two nasals (50% vowel initial, 25% beginning with /m/ and 25% beginning with /n/). Last, half of the words that followed the word-final /t/ had lexical stress at the beginning of the word, in the first syllable, and the other half had lexical stress on the second syllable. An example of each of the conditions for the elicitations, and the resulting sentence is given in Table 1.

Table 1. Three elicitation variables, each with two variations, with an example sentence of each elicitaiton.

Word final-t (verb)	Following word onset	Following word stress	Sentence
vowel_t	**vowel**	non-adjacent lexical stress	a.Teddy ***got exciting*** news from his family.
	consonant		b.Tania ***fought Naomi*** in the semi-finals.
	vowel	adjacent lexical stress	c.Thomas ***got excellent*** news from his manager.
consonant_t	**consonant**		d.Thomas ***got Nina*** a present for her birthday.
	consonant	non-adjacent lexical stress	e.Tucker ***felt nostalgic*** about his college years.
	vowel		f.Tiana ***sent unusual*** things to her sister.
	consonant	adjacent lexical stress	g.Tiana ***sent Nicky*** to get everyone coffee.
	vowel		h.Thomas ***spent everything*** on his renovations.

3.3 Experiment 1

3.3.1 Design

Participants recorded 60 sentences, 48 to be measured and 12 filler sentences. The measured stimuli sentences consisted of 6 verbs in the *vowel_t* and 6 in the *consonant_t* condition. Each verb had 4 iterations: the verb is followed by one of each of the following, a *vowel* or a *consonant*, and that vowel or consonant will have either *adjacent lexical stress*, or *non-adjacent lexical stress*. Each participant received the same sentences, in randomized order.

3.3.2 Procedure

Participants were provided with written instructions for the procedure of the experiment, presented on screen. The experiment and recording began when the participant indicated they were comfortable with the instructions. All slides advanced automatically every 10 seconds. Participants were presented with a contextual sentence in black typeface, and 10 seconds later were presented with a sentence (in dark red typeface) they were to read out loud. The word following the verb was written in bold and italicized typeface: participants were instructed to speak the phrases as naturally as they could, while emphasizing the word in bold and italicized font. They were given 10 seconds to record the sentence and the slide would advance to the next contextual sentence.

3.4 Experiment 2: Production Planning

3.4.1 Design

Participants recorded 21 sentences, 12 to be measured, and 9 filler sentences. The measured stimuli sentences consisted of 6 verbs in the *vowel_t* and 6 in the *consonant_t* condition. Each verb has only one iteration: it's followed by only one instance which meets the following criteria, *vowel* or a *consonant*, with either *adjacent lexical stress*, or *non-adjacent lexical stress*. The conditions were balanced so that there were 6 following words beginning with consonants and 6 with vowels, with an equal balance also of adjacent or non-adjacent lexical stress.

Each participant received the same sentences, but with one of three different speech primers: these included a sentence with the word-final t target missing, a sentence with the following word missing, and one with no words missing. These primer-types were balanced into 3 experiments (Table 2), and these were distributed as close to equally as possible: three participants received experiments 2a, and 2b, and four received 2c. In this experiment, none of the contextual sentences included the t-final verb, as this word was meant to be hidden, as a variable.

Table 2. Three speech prompt variations, and the count of each type in the analysis.

x	freq
full prompt	28
hidden following word	26
hidden word-final t	27

3.4.2 Procedure

Participants were provided with written instructions for the procedure of the experiment, presented on screen. The experiment began when the participant indicated they were comfortable with the instructions. Participants were presented with a contextual sentence in black typeface and were instructed to click next when they were ready to read the sentence for speech (presented in dark red typeface). They were told that if there was a missing word in the sentence they were presented with, they were to wait for the word to appear before speaking the sentence, but in either case, they would have 10 seconds to record the sentence. The word following the verb was written in bold and italicized typeface: participants were again instructed to speak the phrases as naturally as they could, while emphasizing the word in bold and italicized font. After 10 seconds the slide would advance to the next contextual sentence.

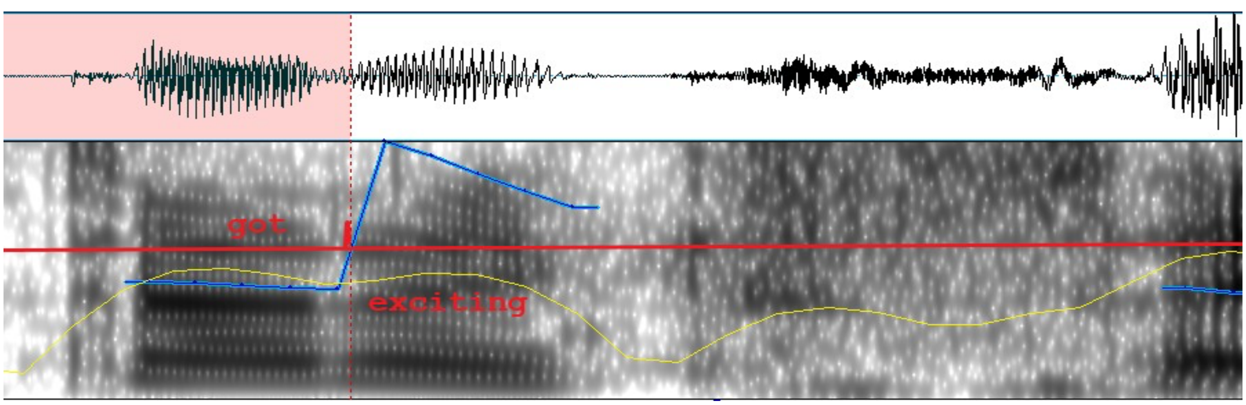
3.5 Measurements

Excel, ELAN (Version 6.4, 2022) and the Montreal Forced Aligner (McAuliffe et al., 2017) were used to annotate the sentence and phonetic transcriptions in Praat textgrids. Aligner transcriptions of the target words were verified and corrected manually where necessary. A manual classification of "T-type" (see discussion in 3.5.1), as well as acoustic measures ('t-word' duration, following word syllable duration, as well as pitch and intensity of both), were annotated and extracted from Praat (Boersma & Weenink, 2022).

3.5.1 T-types, or t-sandhi alternations, annotations

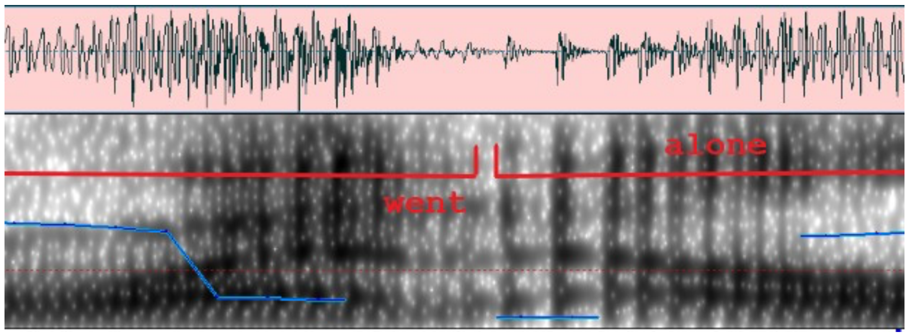
Flaps were annotated as 'fl'. Flaps are most easily distinguishable through listening to the recording but are also identifiable through close proximity of the word-final /t/ word, the onset of the following word (i.e., no long pauses), and a brief closure between the two (Figure 1).

Figure 1. A segment of the words "got exciting", where word-final /t/ is flapped.



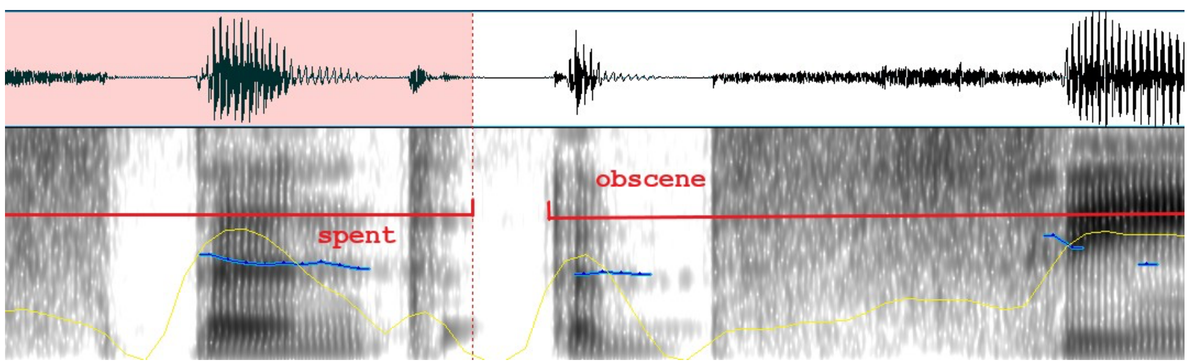
Glottalizations: Glottalizations were annotated as 'gl'. Glottalization occasionally occurred both in the t-final word and the following word, with only a brief closure between the two. Glottalizations were considered a transformation of voicing, and were analyzed as separate from glottal stops (Figure 2).

Figure 2. A segment of the words “went alone”, marked as glottalization in Praat.



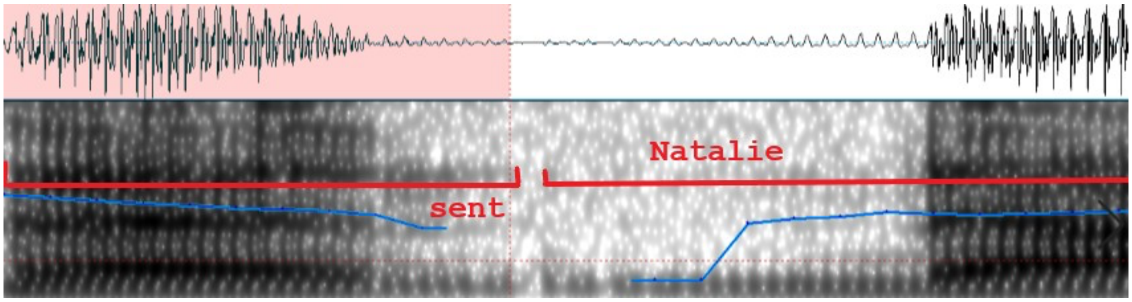
Released t’s were annotated as ‘sr’ for strong release or ‘wr’ for weak release. Strong releases were more clearly followed by aspiration and showed a clear release on the spectrogram while weak releases showed/had weaker bursts, or a less clear release (Figure 3).

Figure 3. A segment of the words “spent obscene”, with word-final t marked as a strong release.



Unreleased-t/t-deletions/Glottal stops: because /t/’s with no clear visible or audible manifestation tend to be ambiguous (without articulatory imaging unreleased /t/ and deleted /t/ would be particularly difficult to differentiate), these were annotated as ‘xt’. In measuring duration of the target words, if there were no discernible silences between the t-final word and the following word, the end of the xt tokens were marked at the nearest 0 crossing of the beginning of consistent voicing of the following word (all following words were voiced) (Figure 4).

Figure 4. A segment of the words “sent Natalie”, marked as glottalization in Praat. Note the lack of burst, glottalization or segments resembling a flap.



3.5.2. Durations and Pause annotations

Any closures between target words that were visible in Praat were marked as ‘pause’ on an interval tier; pauses of less than 0.110s were not annotated. The interval demarcating the /t/-final words were marked as ‘t’, and the following word was split into syllables, and marked as ‘s1, s2, s3, etc.’. These intervals were used to measure duration, pitch and intensity of the /t/-final word and each syllable of the following word.

3.5.3. Stress annotations

Lexical stress that was adjacent to the t (that is, carried on the first syllable of the word) was marked as ‘1’, when the second syllable was stressed, the syllable was marked as ‘2’ to ensure recorded elicitations matched the lexical stress production expected in the experiment design.

4. Analysis

Excel and RStudio were used to analyze the data pulled from Praat textgrids.

4.1 Exclusions

Out of 480 produced tokens in experiment one and 120 in Experiment 2, a total of 180 were excluded for unnatural speech (unusually long pauses between the target words), an unexpected accent, unexpected lexical stress, accidental mispronunciation or misinterpretation of the instructions. Of these exclusions, one participant was omitted completely from the analysis as they were unexpectedly speaking with an accent that was regionally dissimilar from the rest of the group. Another participant was omitted from the results in Experiment 2 as a misinterpretation of the instructions led to unnatural speech.

One participant only produced two out of the four types of t-sandhi: their tokens were omitted from results which looked at proportional t-sandhi distribution, but their data was included for within-participant observations, or within t-sandhi production, like duration, intensity, and pitch. One target following word in Experiment 1 was produced with inconsistent lexical stress and its data was excluded completely. Mispronunciations were omitted even if they were corrected by the speaker as the corrective speech often included over-pronunciations and intonation not produced in natural speech. Strong releases (sr) and weak releases (wr) were only differentiated in the annotation for accuracy as they were fairly distinguishable from each other. As there was no theoretical reason to categorize them separately, they were combined for analysis into one category: r = release.

In Experiment 1, a total of 480 t-sandhi tokens were analyzed: 339 tokens from 9 speakers remained after exclusions. In Experiment 2, 120 tokens were analyzed and 81 remained after exclusions (Table 3), from 7 speakers.

5. Results

5.1. T-sandhi distribution of all participants in both experiments

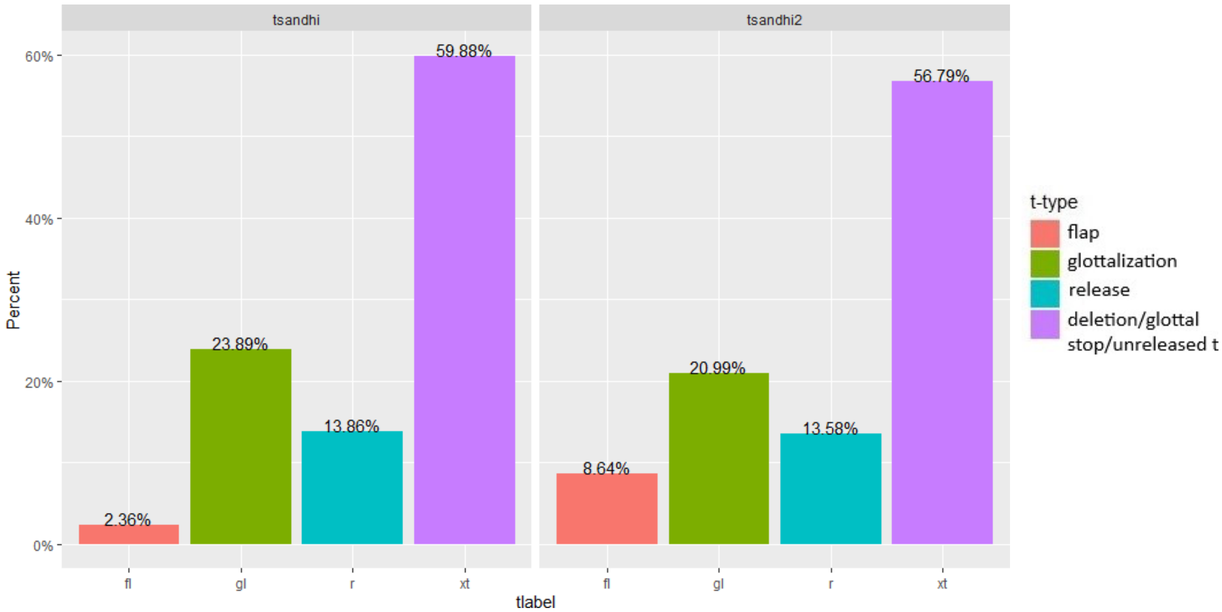
In Experiment 1, of all variations of t-sandhi, the flapping rate was lowest at 2.4%; the rate of released-t was 13.8%; the rate of glottalization was 23.9%, and xt's made up the highest proportion of t-sandhi variation, at 59.9% (Figure 5). In Experiment 2, flapping made up a higher proportion of all t-sandhi, at 8.6%; the rate of released-t was 13.6%; glottalization, 21%; and xt's made up 56.8% of t-sandhi pronunciation.

Overall, across both experiments, among participants who showed all four variations of t-sandhi, the occurrence of flapping was lowest, at only 3.6%, a much lower number than quoted by other studies on flapping. Released t's were produced at a rate of 13.8%, glottalizations at 23.3% and by far the majority of t-sandhi produced were annotated as unreleased-t/t-deletion/glottal stops (xt's) at 59.3% (Table 3).

Table 3. Total number of t-types produced in both experiments.

	tsandhi	tsandhi2	Sum
fl	8	7	15
gl	81	17	98
r	47	11	58
xt	203	46	249
Sum	339	81	420

Figure 5. T-sandhi distribution in Experiment 1 and 2.



One participant produced no releases or flaps in either experiment: instead, they produced a relatively even number of xt and glottalized tokens. The rest of the participants produced all of the possible t-sandhi variants, but flaps were less common in experiment 1 (Figure 6a and b), with only half of participants producing them.

Figure 6a. T-sandhi distribution of all participants in Experiment 1.

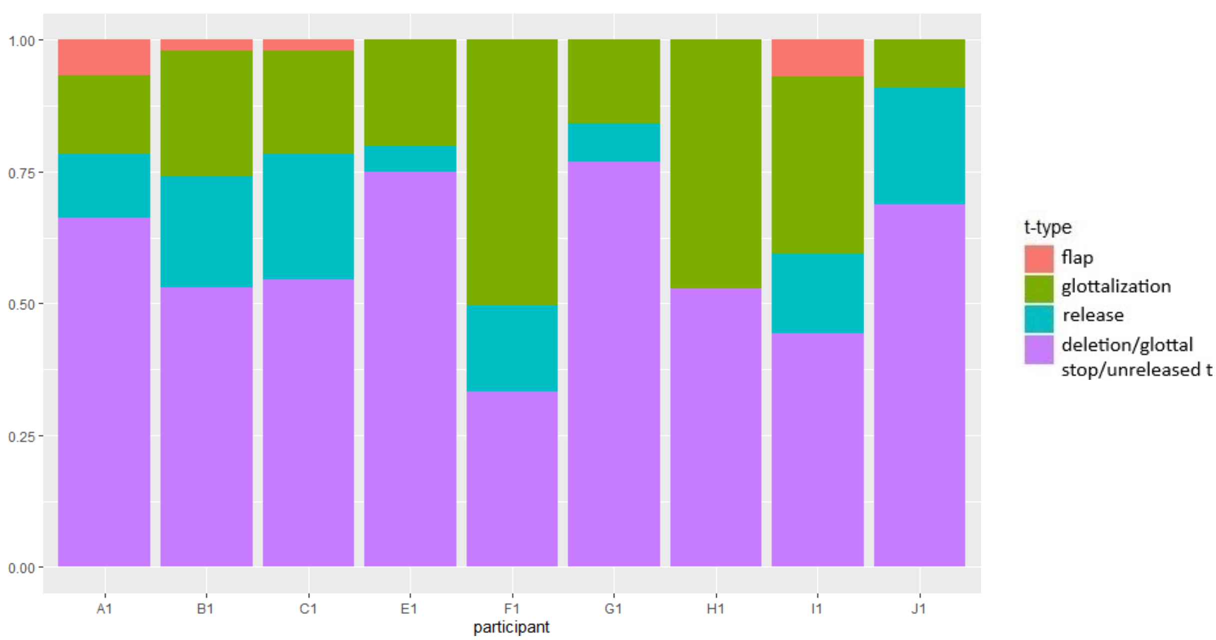
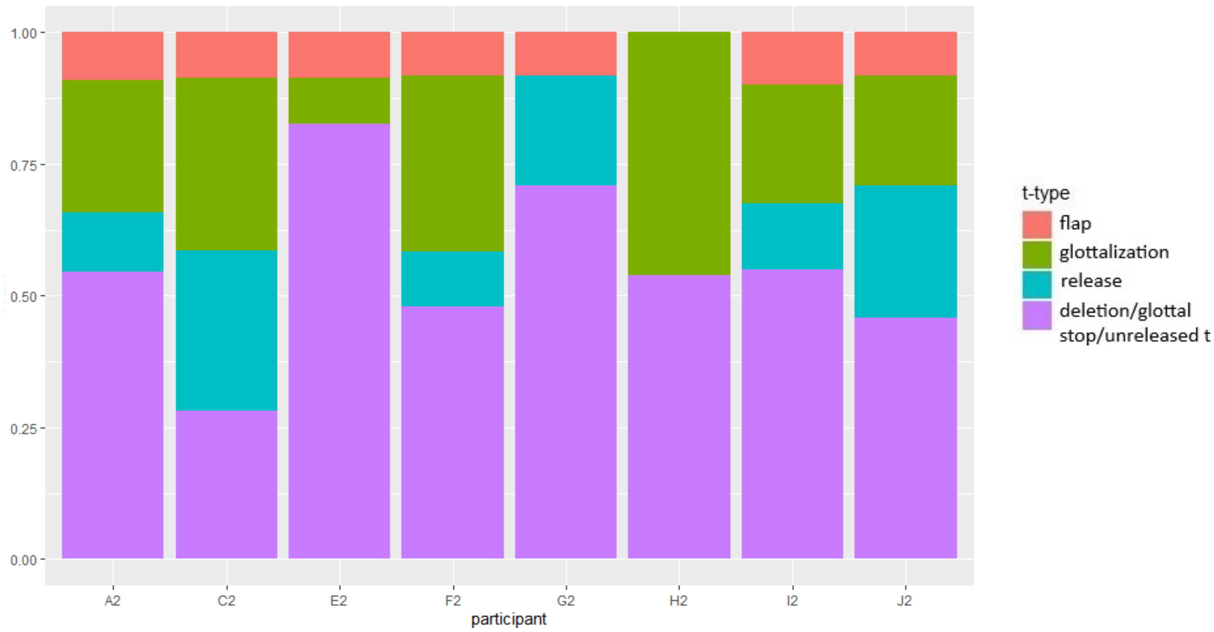


Figure 6b. T-sandhi distribution of all participants in Experiment 2.



5.2 Experiment 1

5.2.1 Segmental phonological conditioning

Since flaps can only occur between two vowels, flapping was expected to only occur in this condition, which was observed in the results. Glottalizations were higher when /t/ was preceded by a vowel instead of a (nasal) consonant (32.6%), and highest in between two vowels (38.6%). Unpronounced /t/, conversely, made up the highest proportion between two consonants and decreased in production before vowels, though they also made up the highest proportion of t-sandhi across all conditions. This proportion was not always very large: xt's were only 4.7% higher than releases in the *vowel_consonant* environment, and only 3% more than glottalized-t in the *vowel_vowel* environment. (Figure 7). The rate of released t's (41%) was highest in the environment where word-final t was preceded by a consonant and followed by a vowel (Figure 7, Table 4) and lowest (at 0%) in the environment where word-final-t was preceded by a vowel and followed by a consonant. Between two vowels, the rate of released-t was 10%, and between two consonants, 3.8%.

Figure 7. The distribution of t-sandhi in the four phonological conditions of Experiment 1.

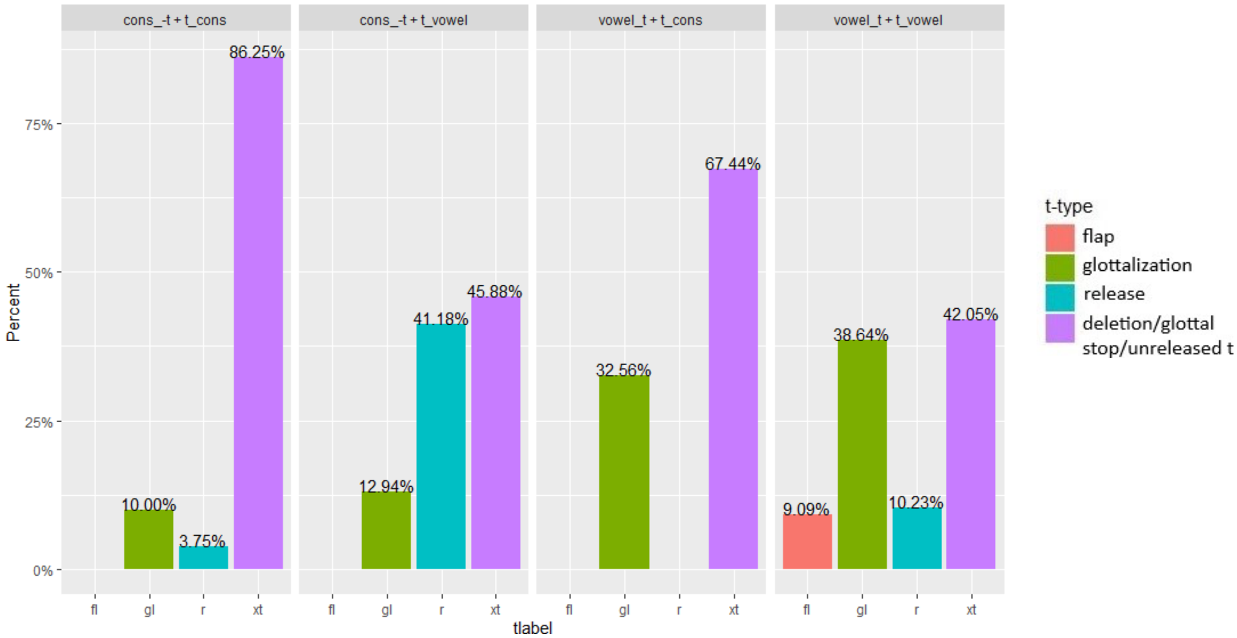


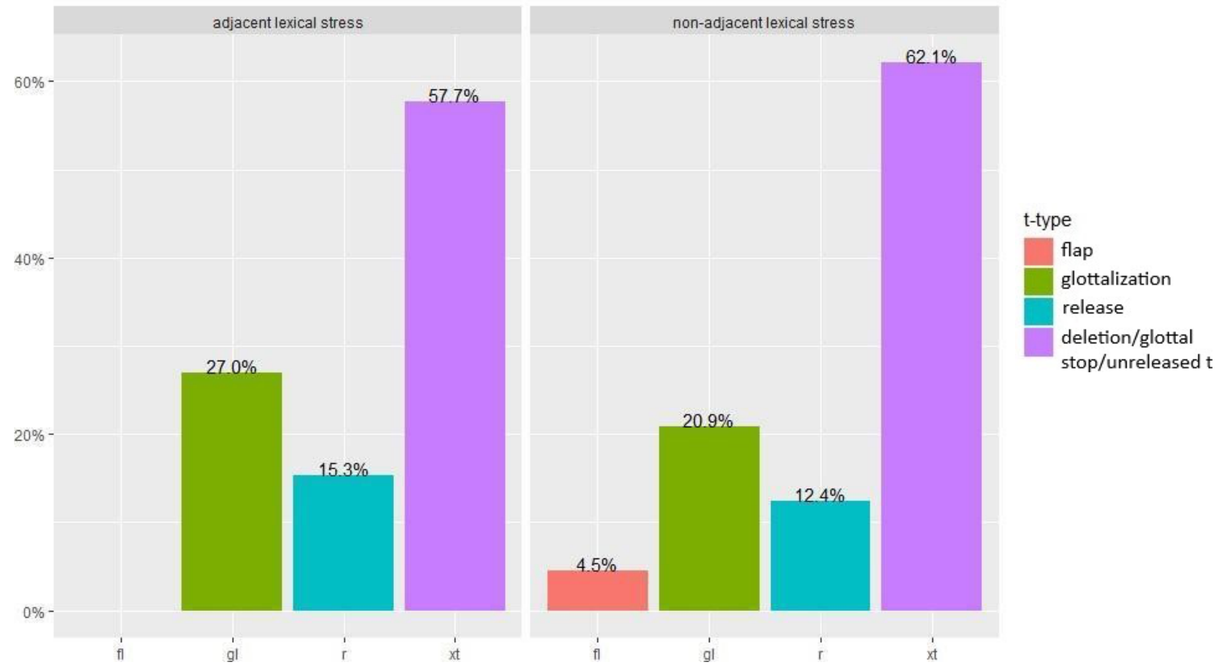
Table 4. T-sandhi tokens across phonological conditions in Experiment 1.

	fl	gl	r	xt	Sum
cons_-t + t_cons	0	8	3	69	80
cons_-t + t_vowel	0	11	35	39	85
vowel_t + t_cons	0	28	0	58	86
vowel_t + t_vowel	8	34	9	37	88
Sum	8	81	47	203	339

5.2.2. Stress conditioning

Stress affected flaps most clearly, with flaps occurring only in environments with non-adjacent lexical stress (Figure 8). The rate of unpronounced-t production also increased where lexical stress was not adjacent. Glottalization and releases did the opposite and increased when /t/ was adjacent to a word with first syllable lexical stress.

Figure 8. The effect of lexical stress adjacency on t-sandhi in Experiment 1.



5.2.3 Combination of effects

The flapping results in Figure 7 were expected (flaps only produced in the *vowel_vowel* condition), but Figure 9 shows the distribution of flaps is even narrower, appearing only when lexical stress on the following word is non-adjacent, that is- when the stress of the following word is not produced on the first syllable: e.g., in Table 1.a: ex**cít**ing (lexical stress in bold and accented): when this condition was met, the rate of flapping was 22.9%.

The released-t results were unexpected: between two consonants, the released-t rate was 3.8% (Figure 7, Table 5): but when lexical stress adjacent, pronunciation rate goes up to 6.98% (0% when lexical stress was non-adjacent, e.g., the string in Table 1.e: “felt nostalgic”); the same reduction in released-t pronunciation (when stress is non-adjacent) is observed between vowels. The results in the overall analysis of the stress effect condition are positive when adjacent to lexical stress: released /t/ appears more likely before a stressed syllable. When combining the segmental condition that is most likely to produce released /t/ (that is after a consonant, before a vowel) with stress, however, the opposite effect is observed: this rate jumps from 35.6% to 47.5%.

The number of ‘*xt*’ pronunciations before consonants remains highest regardless of the stress condition, but it also increased when the adjacent syllable was not stressed, regardless of the segmental context.

The opposite was found with glottalization, where most segmental contexts see an increase in glottalization production when the following segment is stressed, the exception to this was between two consonants, which seemed to increase glottalization when the following segment was unstressed: a result which is opposite to the overall effect of stress on glottalization rate.

Figure 9. T-sandhi distribution, Experiment 1: segmental condition by stress.

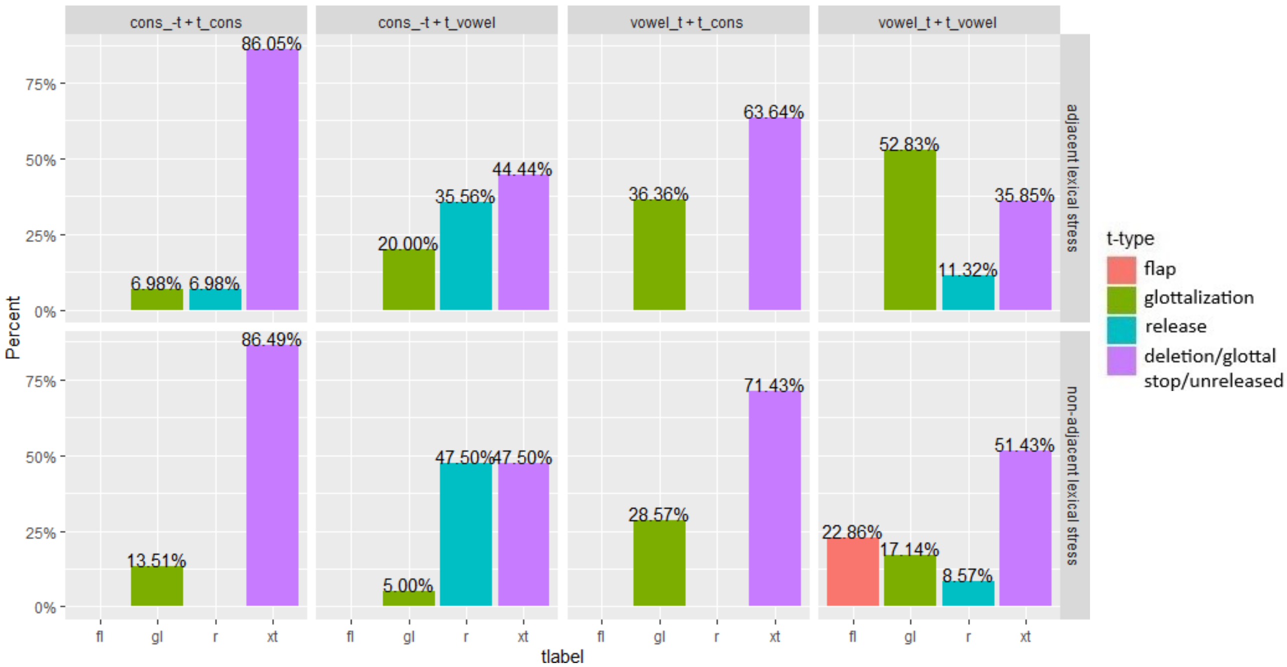


Table 5. Tokens produced in each of the phonological/prosodic conditions in Experiment 1.

	adjacent lexical stress	non-adjacent lexical stress	Sum
cons_-t + t_cons	43	37	80
cons_-t + t_vowel	45	40	85
vowel_t + t_cons	44	42	86
vowel_t + t_vowel	53	35	88
Sum	185	154	339

5.3 Experiment 2

5.3.1 Production Planning and segmental conditioning

The aim of this experiment was to be able to compare the results of the full prompt condition (i.e., stimuli where the participant did not experience blanks in the first visual speaking prompt) to those where either the /t/-final word (i.e., the verb) or the word following the /t/-final word were initially hidden. The results show that delaying the word following the /t/ has the largest effect on t-sandhi production: flapping rates are reduced and glottalization increases, decreasing the number of xt pronunciations (Figure 10).

Delaying the availability of the word following /t/ seemed to inhibit flapping the most: among tokens available for flapping (based on the segmental *vowel_vowel* condition), 75% were flapped in the full prompt condition, 25% in the hidden following word condition, and 50% in the hidden t-final word condition (Tables 6, 7, Figure 11).

Releases seemed to be affected by this condition as well: in the segmental condition where releases are produced in the highest proportion (*consonant_vowel*), delaying the availability of the following word reduced released-t pronunciation. Unexpectedly, delaying access to the word-final t increased releases (Figure 11). As in Experiment 1 (see Figure 10), the *vowel_consonant* condition produced no released t's, regardless of prompt type, but the other two conditions (*consonant_consonant*, and *vowel_vowel*) showed opposite directions of released-t production, based on prompt type.

Glottalization rates remained the same when the verb was hidden, however hiding the upcoming word increased glottalization (Figure 10).

Figure 10. T-sandhi distribution by prompt type, Experiment 2.

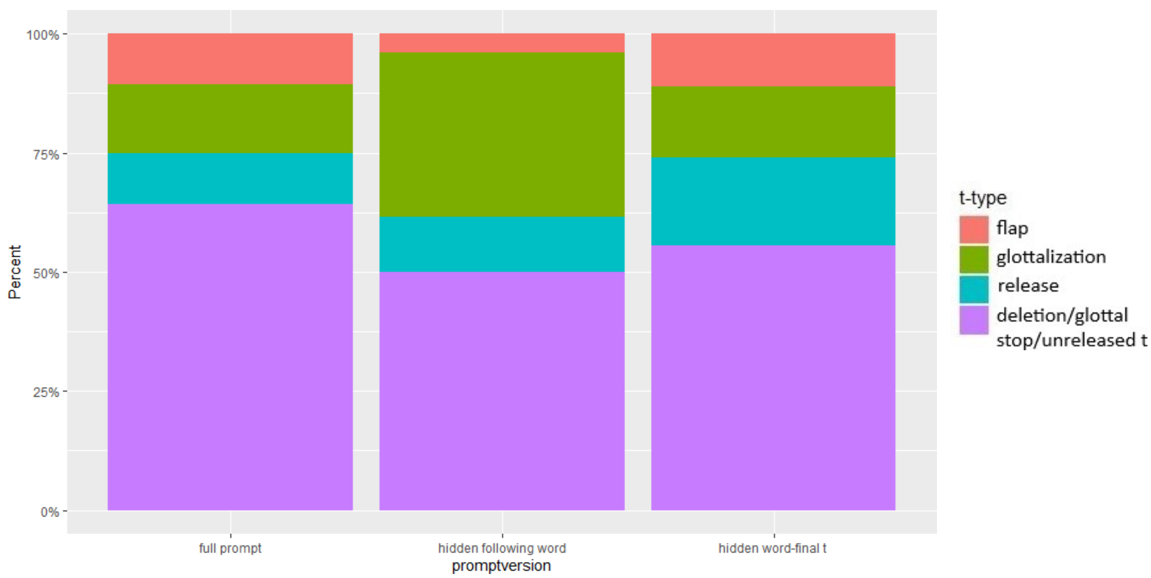


Figure 11. T-sandhi production by segmental context and prompt type, in Experiment 2.

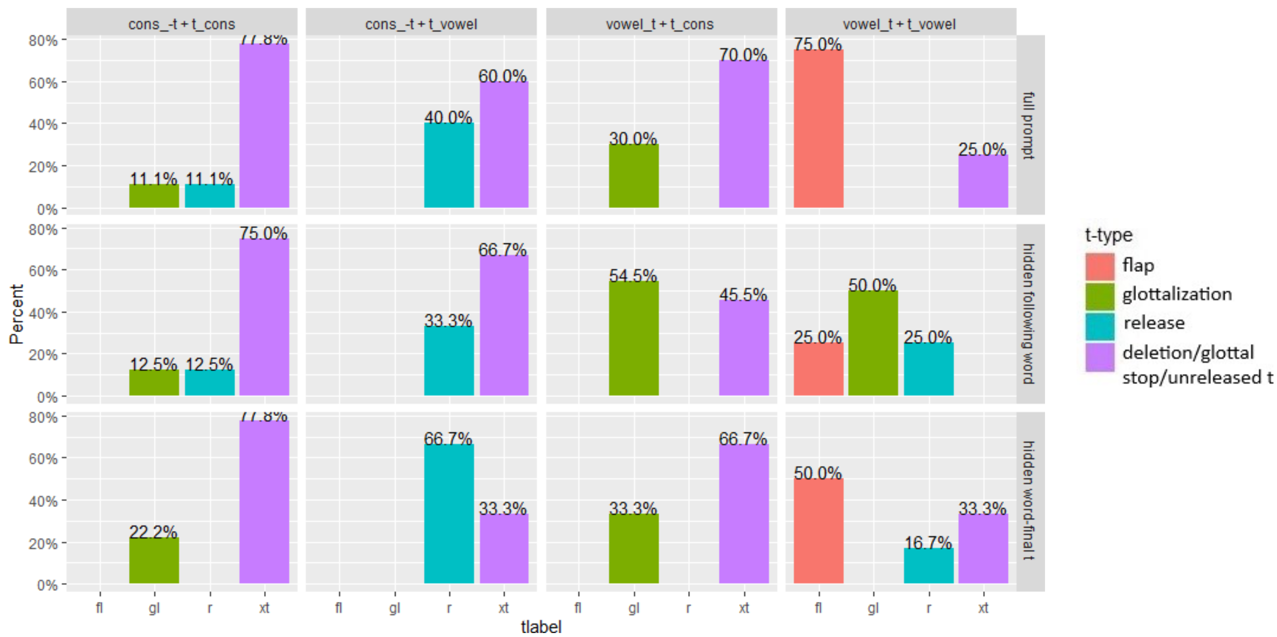


Table 6. Tokens analyzed in Experiment 2, grouped by prompt type and segmental environment.

	full prompt	hidden following word	hidden word-final t	Sum
cons_-t + t_cons	9	8	9	26
cons_-t + t_vowel	5	3	6	14
vowel_t + t_cons	10	11	6	27
vowel_t + t_vowel	4	4	6	14
Sum	28	26	27	320

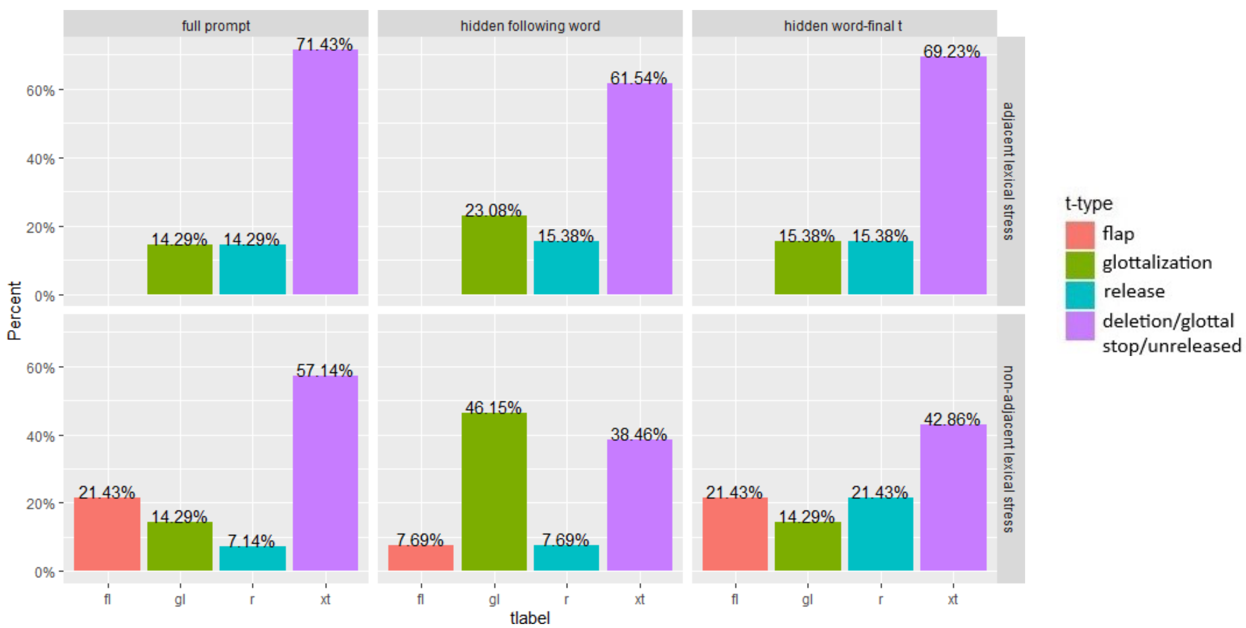
Table 7. T-sandhi tokens, in Experiment 2, grouped by prompt type.

	full prompt	hidden following word	hidden word-final t	Sum
fl	3	1	3	7
gl	4	9	4	17
r	3	3	5	11
xt	18	13	15	46
Sum	28	26	27	81

5.3.2 Production Planning and stress

The overall distribution of t-sandhi in each of the prompt versions was compared to the stress conditions in Figure 13. Ignoring the flapping data, as the data showed a clear stress condition requirement, all stress conditions show distributions that are close to their corresponding overall results in Figure 10 (as in the previous section), though the results of non-adjacent stress more clearly match those distributions, suggesting that following stress might be a stronger effect here. As Experiment 2 had a much smaller pool of tokens, these results are less likely to be robust.

Figure 13. T-sandhi distribution by prompt type, and adjacent lexical stress, Experiment 2.



5.4 Boundary Strength

5.4.1 Duration of the t-word

The duration of flapped words was shortest, and the duration of released t's was longest, while unpronounced t's and glottalizations fell in between, a durational trend which is common across all effects. This difference in duration was significant for flapped words and released t's (Figure 15), which is also common across all effects. As mentioned in the segmental effects results, releases occurred much more frequently when part of a complex coda: the duration of verbs with complex codas was significantly longer when the t was released (Figure 16). In the simple coda condition (t following a vowel), the duration of flapped words was significantly shorter.

T-sandhi durations were much more distinct when followed by a vowel, even showing a rare difference between the durations of xt's and glottalized /t/, though this difference was not significant (Figure 17). They were also more distinct in the non-adjacent stress condition (Figure 18).

In terms of the effects of the segmental environment, it seems that the preceding condition (complex coda, vs vowel only) has a lengthening effect on the word duration of all types[3] of t-sandhi: in releases and unpronounced /t/ (Figure 19), this effect is significant. Being followed by a consonant (Figure 20) lengthens the duration of t-final words pronounced with releases and glottalizations, but reduces them before unpronounced /t/, however this effect is not significant. The stress adjacency condition (Figure 21) shows the same grouping of t-sandhi effects: non-adjacent stress lengthened the duration of t-final words pronounced with glottalized or released /t/, and shortened the duration of unpronounced /t/ words, though none of these differences were significant either.

Figure 15 . T-sandhi distribution by duration of the (word-final t) verb across both experiments.

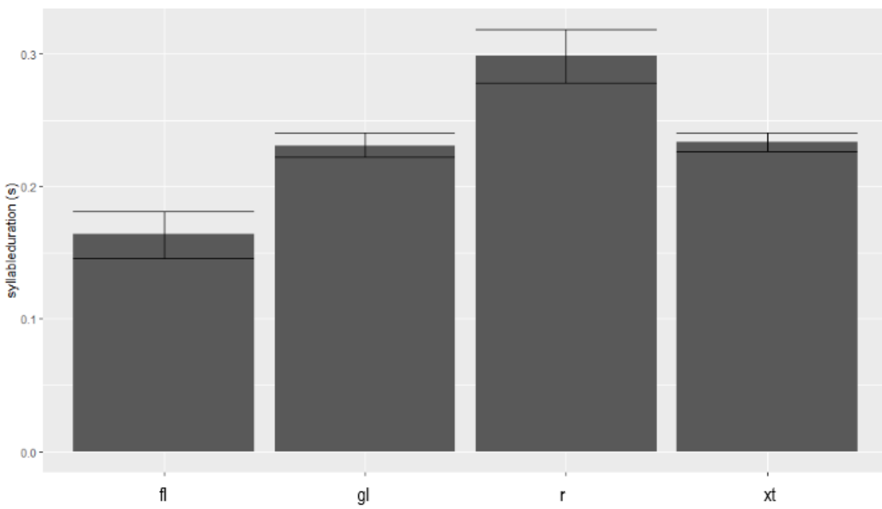
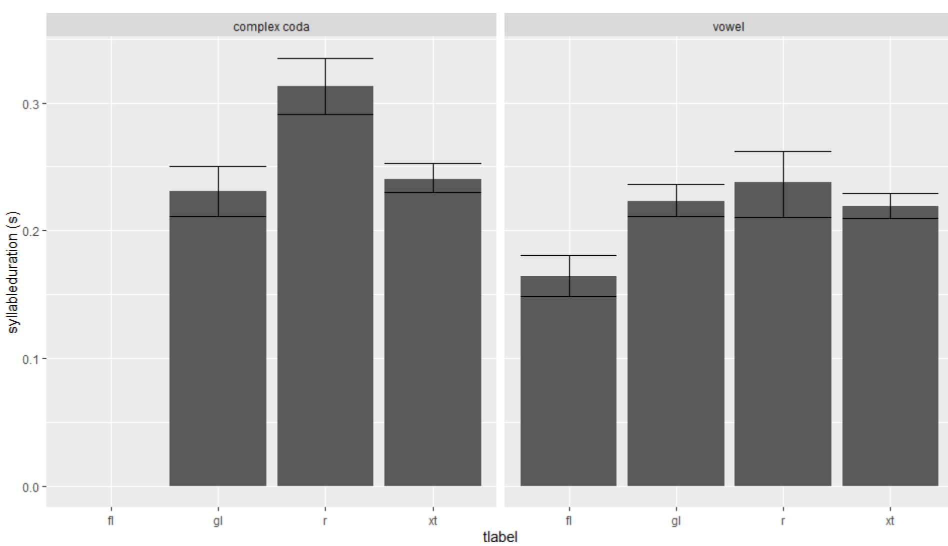


Figure 16. Duration of the verb and t-sandhi variation, grouped by preceding context.



---

[3] Segmental and prosodic effects on flaps are ignored as their appearance requires specific environments.

Figure 17. Duration of the verb and t-sandhi variation, grouped by following consonant or vowel.

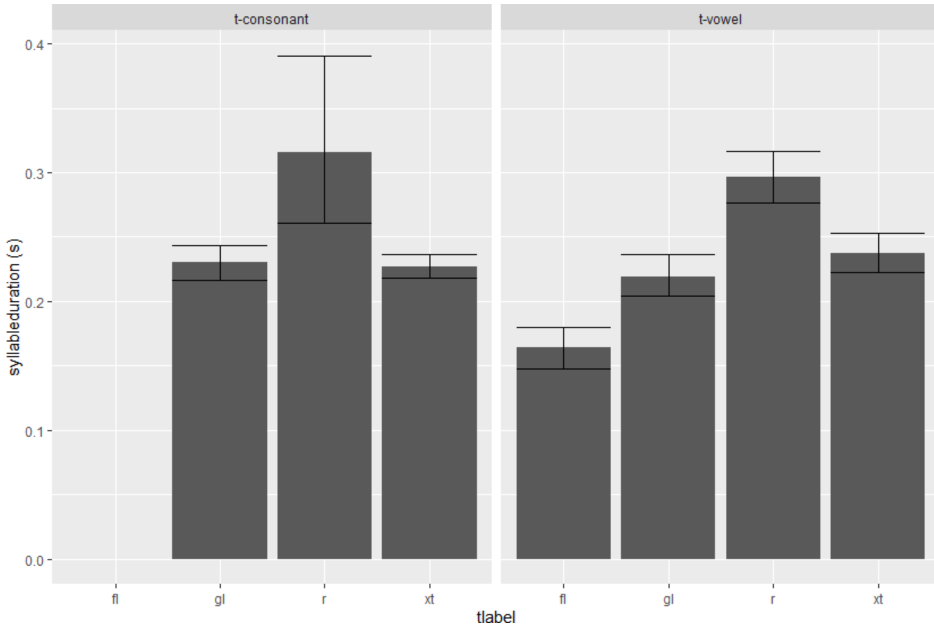


Figure 18. Duration of the verb and t-sandhi variation, grouped by the following stress condition.

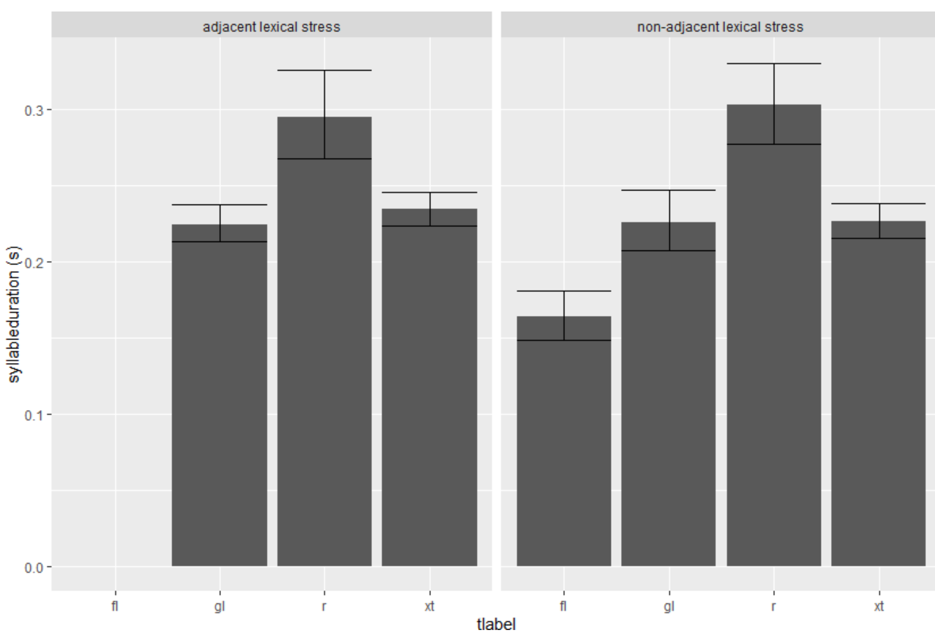


Figure 19. Effect of preceding environment on the duration of the verb (grouped by t-sandhi variation)

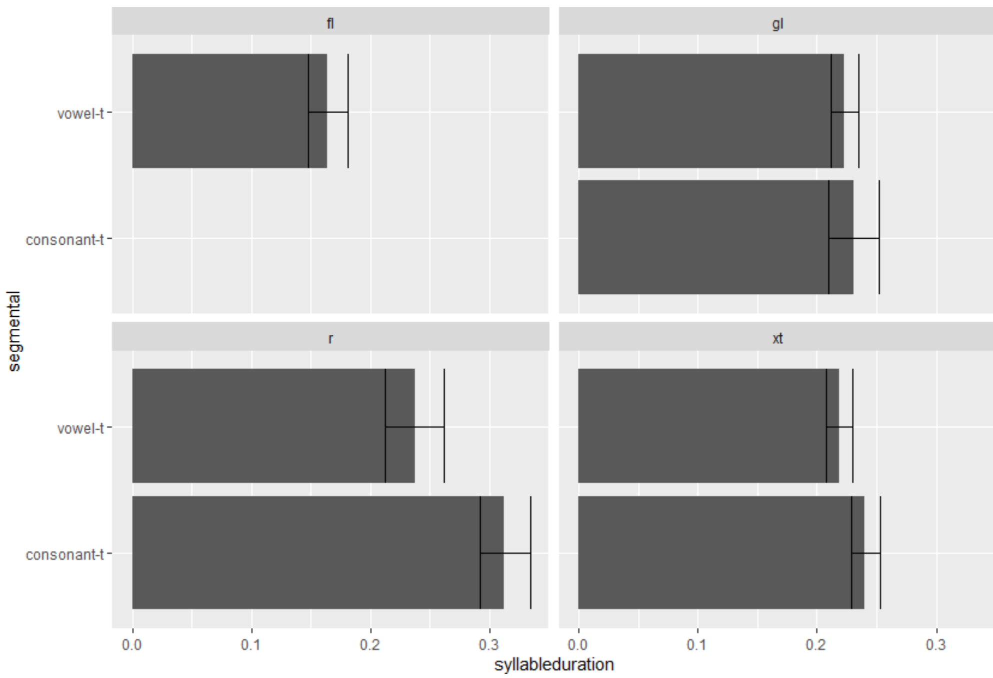


Figure 20. Effect of following environment on the duration of the verb (grouped by t-sandhi variation)

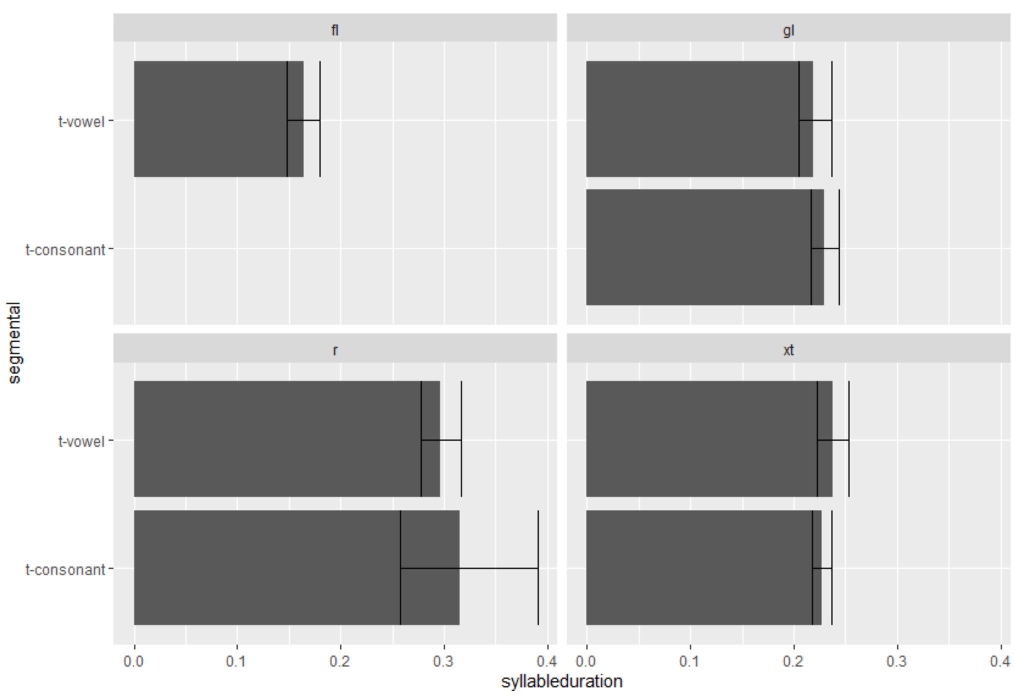
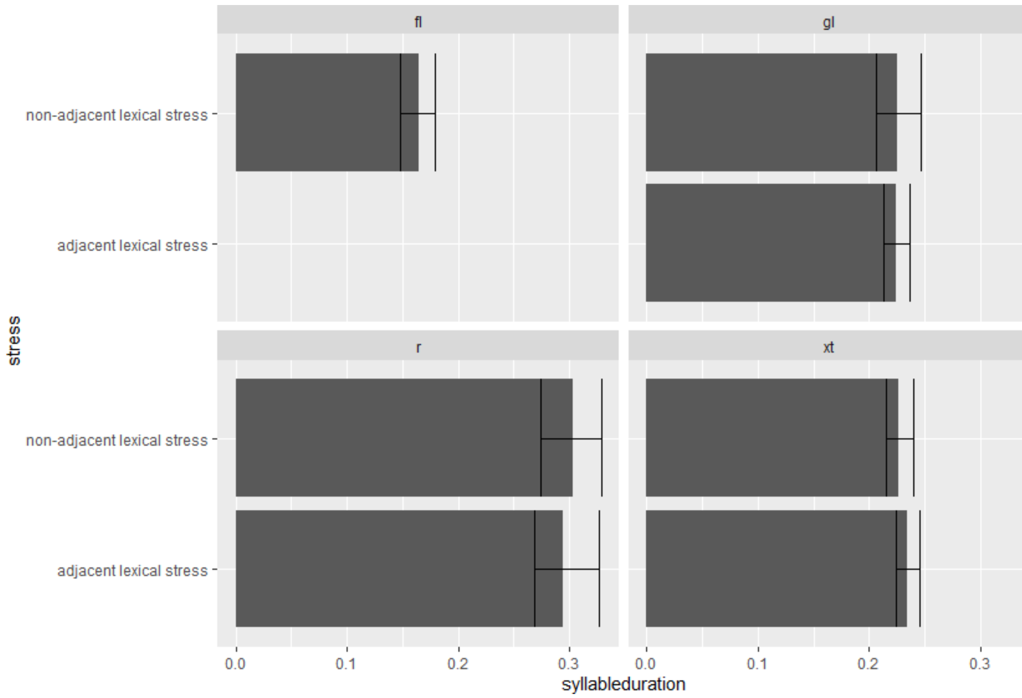


Figure 21. Effect of following stress on the duration of the verb (grouped by t-sandhi variation)



5.4.2 Pauses

Because unpronounced /t/s have boundaries that are difficult to separate from any pause that might occur before the following word, the duration of the verb was combined with any following pauses to see if it influenced the distribution of the duration of t-sandhi rates (Figure 18). The results were as identically significant as those in Figure 18.

The number of pauses that followed each t-sandhi variation were counted. No pauses occurred with flaps, and few occurred with unpronounced /t/ tokens. More than half of released and glottalized tokens included a pause.

Figure 18. T-sandhi distribution by duration of both the (word-final t) verb and any following pauses.

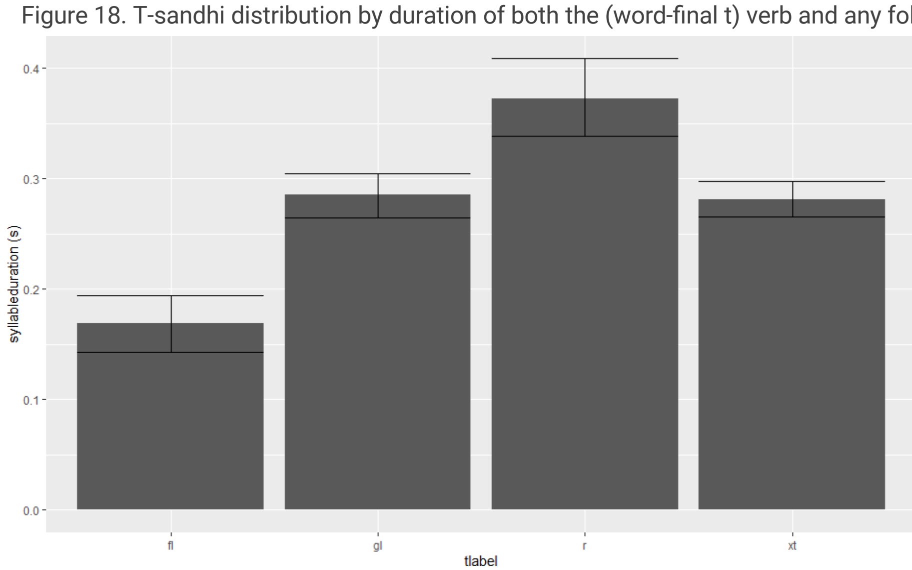
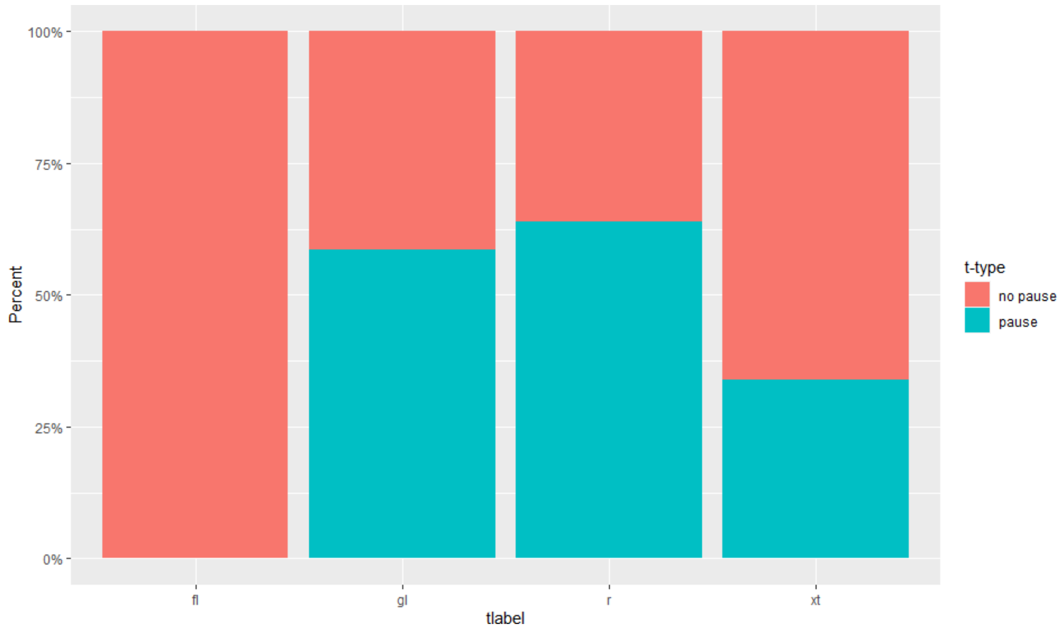


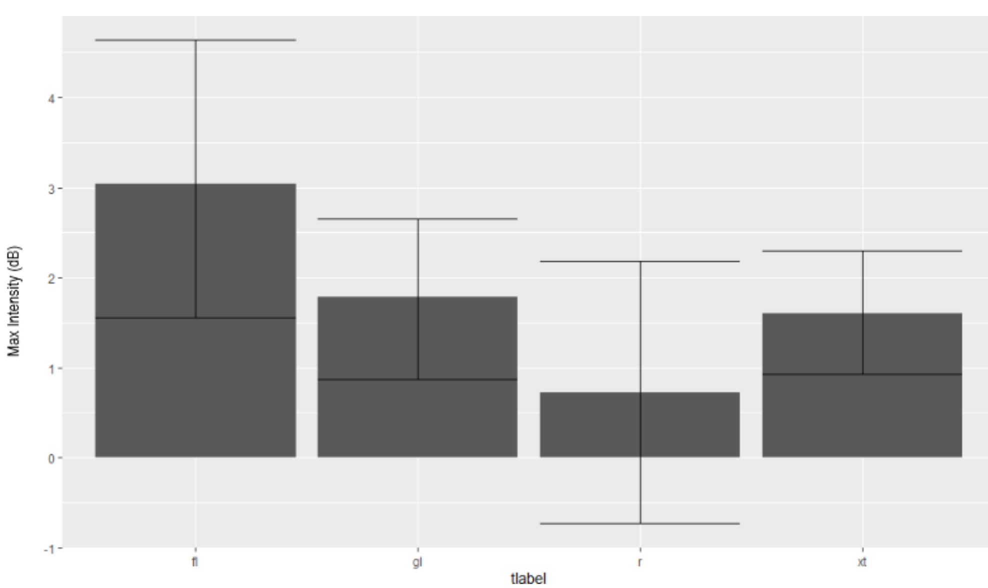
Figure 19. Percentage of pauses between word-final t and the following word, by t-variation.



5.4.3 Intensity and Pitch

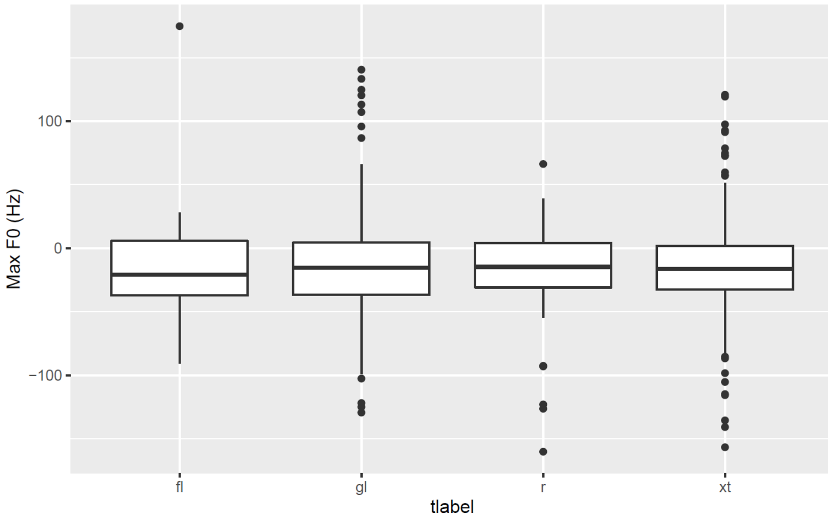
There were distinguishable (though insignificant) effects on the difference in intensity between word-final t and the following syllable (Figure 20): flaps showed a much higher difference between the two segments and released /t/ showed a much lower change in intensity between the two. Intensity differences between the first syllable and glottalized or unpronounced /t/'s fell in between flaps and releases and were somewhat equal.

Figure 20. The delta of the intensity between the following syllable, and the (word-final t) verb.



Pitch results were more difficult to analyze, as there were outliers in the data that made trends more difficult to observe. It does appear that the pitch delta between the verb and the following syllable (Figure 21) is larger and more varied for glottalized and xt tokens, and smaller for flap and released tokens, without significance.

Figure 21. Boxplot of the pitch delta between the following syllable, and the (word-final t) verb.



6. Discussion

6.1 T-sandhi distribution

Since there were no studies that seemed to explore the overall distribution of t-sandhi there were no predictions made: perfectly balanced t-sandhi production was not anticipated but the strength of the effect of each type of phonological conditioning in this particularly narrow prosodic environment was unknown; the large proportion of glottalized and unpronounced-t tokens observed in the data was an interesting result. Studies in t-sandhi tend to frame a particular variation of /t/ pronunciation as a binary variant of either an unnamed or assumed released t-default. In the case of word-final CSD in particular, finding the predictability of an unpronounced-t seems almost impracticable if it’s the most common form of t-pronunciation. The phonological motivation (and thus the predictability of) t-deletions, for example, are different from glottal stop production, but the auditory unintelligibility between deleted /t/, glottal stops, and unreleased /t/, adds a layer of ambiguity which prevents easy confirmation of their presence in a suspected environment (when relying on acoustic measures alone). With more analysis it might be possible that other acoustic correlates, perhaps ones that are also measures of boundary strength, might provide ways to disambiguate these three ‘unpronounced’ forms from one another.

The within participant results were also interesting, as even though the speakers were native speakers of the same language in the same region, it was clear that not all speakers used the same repertoire of t-allophones in their natural speech. This had interesting implications for studies of single allophonic variation, e.g., one which only explores flapping, as including participants who never flap skews results. Excluding these participants might be less important in corpus studies, which involve a much larger set of data and a less exploratory set of conditioning, but would be important in any production study.

The production rate of flapping was interesting in a number of ways: at 3.6% across both experiments, this rate was much lower than the 22.01 % found in the Kilbourne et al. study (2016), or the 93.9% rate found in the Patterson and Connine study (2001). The higher rates in these studies might be the result of imbalanced data: these studies, like the current one, used intervocalic-t as the environment that could produce flapping, but based on the results of Experiments 1 and 2 in the current study, specifying an unstressed adjacent syllable is also a necessity. If a large proportion of the vowel-initial words that followed word-final t met this condition in the corpus, the proportion of flapping would be much higher. And if a control for minimum word size, or at minimum a prosodic word, was not introduced, including bigrams with the indefinite determiner ‘a’ is likely to increase the overall flapping production rate. While this might not particularly interfere with the ability of the PPH to predict flapping PPH, as ‘a’ also classifies as a high frequency word, it’s still worth considering that its intrinsic prosodic quality also facilitates flapping. As mentioned in the results, if the rate of flapping in the current study was narrowed to its production rate in stimuli that were intervocalic with non-adjacent stress the flapping rate rises to 22.9%, a rate similar to that found in the Kilbourn et al. (2016) study.

There was no baseline for a prediction of released-t production rate, but this number was lower than expected for a pronunciation which is considered the underlying form of /t/. Their relative rarity, however, was explained by an unexpected segmental condition: released t’s were never pronounced when preceded by a vowel and followed by a consonant (which accounted for 25.3% of stimuli in Experiment 1). Production rate differed between the environments in which /t/ was pronounced as released: in the *consonant_vowel* condition, it was highest, 41%; in the *vowel_vowel* condition: 10%; and finally, in *consonant_consonant*: 3.8%.

Though duration was used as an indicator of boundary strength, the distribution of t-sandhi and the effects of segmental contexts on this duration, were interesting: the duration of glottalized and unpronounced /t/’s were always similar, but the direction of effects (longer before consonants, or non-adjacent stress) was the same for glottalized /t/ and released /t/.

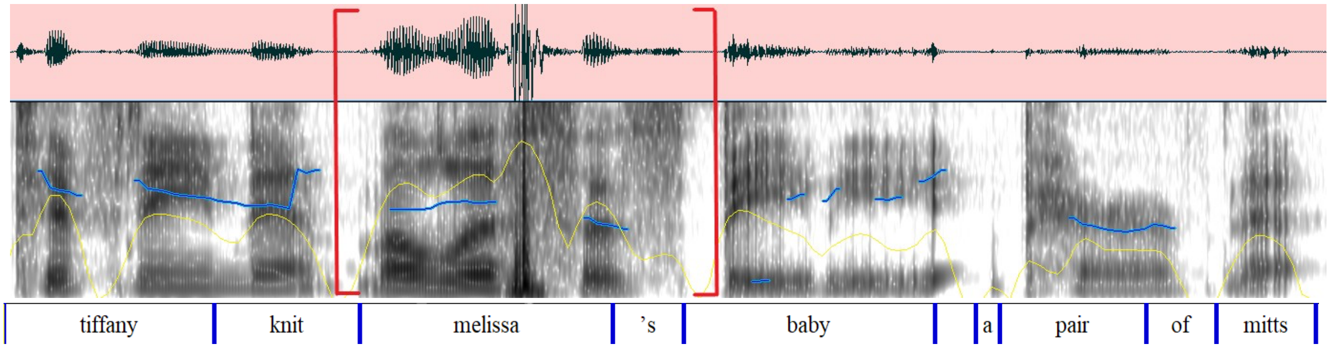
6.2 Stress and boundary strength

At least some of the released-t segmental conditioning may be prosodic: though all t-final words chosen for this experiment were monosyllabic, words with complex codas were longer in duration (Figure 16); of these words, those with released t’s were significantly longer in duration. As studies have shown that increased word duration is associated with a stronger following prosodic boundary (Elfner, 2016), these results suggest that released-t production might be promoted by stronger boundaries. This interpretation is strengthened by the effect of lexical stress adjacency (Figure 8) which shows that released-t rates are higher when the following word's first syllable is stressed. There are some qualities of released-t that might explain a preference for stronger boundaries: consonant aspiration increases the prominence of a syllable (Fery, 2016: 17), and released t's are more common in the onset of a syllable; resyllabification of a released-t pronounced before a weak boundary is more prone to being both perceived as having stress, or being produced with (ungrammatical, or incorrect) stress. A strong boundary prevents this by increasing the distance between two words, something which is already known to affect t-sandhi.

Lexical stress adjacency also had a strong effect on flapping. At first glance it might seem that the environment which produces flapping in this study is nothing more than an extension of the condition which sees intervocalic t's categorically pronounced as flaps before unstressed vowels (word internally), however, external t-sandhi flapping was only categorical before unstressed vowels in Experiment 2, indicating there are definitely other effects. Indeed, categorizing the first syllable of any of the following words as unstressed would be inaccurate, as controls ensured that the following word would always be 'stressed' through the participant directed emphasis, even if the lexical stress of the following word was non-adjacent (see Figure 20). Another approach would be to analyze these results as an effect of weaker boundaries between t and the following word, which is supported by the significantly shorter duration of flapped words and though pauses are not required to identify a boundary, pauses are associated with stronger boundaries, and none of the flapping tokens were followed by a pause in this data.

The intensity delta measured between /t/ and its adjacent syllable provided more evidence for the effect of boundary strength on t-sandhi: with flapping, the delta between the intensity of the single syllable final-t word and the following syllable were high, and with released /t/ this delta was low. This seems to show that where the intensity clearly marks a boundary, flapping is a possibility, and where this delta is low, released /t/ production is either serving to increase the boundary between the two, or is able to be produced as a released /t/ as the lowered delta provides more room for tense production, without changing the lexical stress assignment of the following word. This might also explain the availability of flapping, as flapping doesn't serve to increase stress. The pitch data was expected to produce results that had a similar distribution to those in intensity, not particularly with a smaller pitch delta being associated with releases (and a larger delta with flaps), but with clear differences between the t-sandhi types. The results did hint at some associations: these results are intriguing and the implications they have for boundary strength effects make them a promising area for future research.

Figure 20. A spectrogram of the sentence, Tiffany knit Melissa's baby a pair of mitts, with the focused and emphasized word bracketed in red. Even the first syllable of Melissa's is more intense (measured in yellow) than other words in the phrase.



6.3 Production Planning

Of the three prompting conditions in Experiment two, the "hidden following word" condition (in Figures 10-13), which delayed access to the upcoming (following) word, was expected to affect t-sandhi variability the most; as flapping across a word boundary would require access to both words, this was the t-sandhi variation which was predicted to be most affected in this condition. The results seemed to show that delaying the word following the /t/ did have a larger effect on t-sandhi production than delaying access to the word containing the /t/: flapping rates experienced the largest reduction in this condition, from 75% to 25% (and only 75% to 50% in the hidden t-word condition); however, after token exclusions were performed, the conditions in the flapping data were left unbalanced: every token that was eligible for flapping was flapped. If it is assumed that the context of flapping found in Experiment 1 is a requirement of flapping (intervocalic and not followed by adjacent lexical stress) the results in Experiment 2 are less compelling for the effect of the PPH on flapping: what looked like a prompt type effect could actually just be a reflection of vowel type (all flapping tokens had the same vowel in the verb), or coda complexity (none of the verbs had a complex coda).

Other t-sandhi results in the production planning experiment supported the PPH's predictions as well, though not straightforwardly. In Figure 11, the rate of released-t production in the "*consonant_vowel*" environment was reduced when the upcoming word was delayed. However, blocking of the word with the word-final t seemed to increase the production of releases. This could be the effect of a small sample size, or it could be because unlike flaps, which require resyllabification, released t's are not ungrammatical in any position. Since locality is a requirement of the PPH, including this condition might not have yielded results predictable by the PPH.

7. Conclusion

It was proposed that examining the distribution of all types of external t-sandhi produced clause-internally would provide a holistic look at how speakers use t-sandhi distribution in natural speech. The results showed that flaps and releases, the more marked forms of t-sandhi, (in terms of production rates) favored prosodically defined environments not examined in previous studies: word-external, clause-internal flapping exclusively preferred environments where the following word did not carry lexical stress on the first syllable, showing a potential preference for weak prosodic boundaries with the following word. Released-t's exclusively dispreferred environments where /t/ was preceded by a vowel and followed by a consonant, but also seemed to prefer environments where prosodic boundaries were stronger. Studies of released /t/'s seem to remain under-represented and the effects observed here showed the possibility that there are clear segmental and prosodic conditions that affect its production rate, by focusing on those criteria that seemed to produce an abundance of released /t/, there might be even more yet to be observed effects, or clearer motivations behind effects like its seeming preference for stronger boundaries between words, and lower intensity deltas between it and the following syllable. The production planning experiment results seemed to show effects consistent with the PPH, but could benefit from an experiment which focuses exclusively on the effect of upcoming word availability on processes that affect flapping, releases, or both. Based on the results of this study, which showed two fairly distinct distributions, focusing on released and flapping variations of t-sandhi might be able to increase the robustness of the PPH effect, but even simply ensuring that the token availability reflects both controlled segmental and prosodic variables for one effect would be an advantage over the current study. Another area of interest would be the often similar prosodic distribution between glottalization and unpronounced /t/: they often showed similar qualities, e.g., in duration, in intensity delta, but different directions of effects: increasing or decreasing in opposite directions.

Controlling for environments which significantly reduce flapping and released /t/ pronunciations would provide a more in-depth look at the unpronounced /t/ category, and maybe some insight into how these might be disambiguated using acoustic measures. Are the unpronounced tokens that behave like glottalized tokens instances of glottal stops? If a set of effects cause xt-type pronunciations to move in the same direction of released /t/, is unpronounced /t/ more likely to be occurring? This paper fits into the expansive study of t-sandhi by illustrating the value of further examination of the interaction between prosody and external t-sandhi, the benefits of research that considers variations of t-sandhi in tandem, as well as the role of the Production Planning Hypothesis in all of those elements.

References

Boersma, P. & Weenink, D. (2022). Praat: doing phonetics by computer [Computer program]. Version 6.2.14, retrieved 24 May 2022 from http://www.praat.org/

Bybee, J. and Scheibman, J. The Effect of Usage on Degrees of Constituency: The Reduction of Don't in English. (1999). Linguistics, vol. 37, no. 4, July 575–596, doi:10.1515/ling.37.4.575.

Cho, T. and Keating, P. (2001). Articulatory strengthening at the onset of prosodic domains in Korean, Journal of Phonetics 28: 155-190

Coetzee, A.W., Kawahara, S. (2013). Frequency biases in phonological variation. Natural Language & Linguist Theory 31, 47–89. https://doi.org/10.1007/s11049-012-9179-z

Eddington, D. and Taylor, M. (2009). T-Glottalization In American English. American Speech 84: 298-314.

ELAN (Version 6.4) [Computer software]. (2022). Nijmegen: Max Planck Institute for Psycholinguistics, The Language Archive. Retrieved from https://archive.mpi.nl/tla/elan

Elfner, E. (2016). Subject/Object complexity and prosodic boundary strength in Irish. 178-182. 10.21437/SpeechProsody.

Féry, C. (2016). Intonation and Prosodic Structure. Cambridge: Cambridge University Press. doi:10.1017/9781139022064.002

Fougeron, C. and Keating, P. A. (1997). Articulatory strengthening at edges of prosodic domains, J. Acoust. Soc. Am. 101, 3728-3740

Guy, G. (1980) Variation in the group and the individual: the case of final stop deletion. In W. Labov (Ed.), Locating Language in Time and Space (pp. 1-36). Academic Press.

Hoard, J.E. (1971) Aspiration, Tenseness, and Syllabication in English." Language, vol. 47, no. 1, 133–40, https://doi.org/10.2307/412191.

Lehiste, I. (1973). Phonetic Disambiguation of Syntactic Ambiguity. Journal of the Acoustical Society of America, 53, 380-380.

Keating, P., Cho, T., Fougeron, C. and Hsu, C.(2003). Domain-initial articulatory strengthening in four languages. Phonetic Interpretation (Papers in Laboratory Phonology 6), edited J. Local, R. Ogden, R. Temple, Cambridge University Press, 143-161

Kilbourn-Cerón, O., Wagner, M and Clayards, M. (2016) The effect of production planning locality on external sandhi: a study in /t/. In: The proceedings of the 52nd Meeting of the Chicago Linguistics Society. URL: http: //ling.auf.net/lingbuzz/003119.

McAuliffe, M., Socolof, M., Mihuc, S., Wagner, M., Sonderegger M. (2017). Montreal Forced Aligner: trainable text-speech alignment using Kaldi. In Proceedings of the 18th Conference of the International Speech Communication Association

Patterson, D. and Connine, C.M. (2001). Variant Frequency in Flap Production" Phonetica, vol. 58, no. 4, 254-275. https://doi.org/10.1159/000046178

Schmitz D, Baer-Henney D, Plag I. (2021). The duration of word-final /s/ differs across morphological categories in English: evidence from pseudowords”. Phonetica. Oct 25;78(5-6):571-616. doi: 10.1515/phon-2021-2013. PMID: 34699697

Tanner, J. & Sonderegger, M. & Wagner, M. (2017). Production planning and coronal stop deletion in spontaneous speech, Laboratory Phonology 8(1), 15. doi: https://doi.org/10.5334/labphon.96