

English coronal stop deletion is categorical not gradient

Yunting Gu & Ryan Peters*

Abstract. English Coronal Stop Deletion (CSD) has been a subject of debate in terms of whether it is categorical or gradient. Previous studies have overlooked the possibility that tongue tip raising during inaudible coronal stop may come from neutral tongue tip position rather than gradient CSD. The current study found that in sentence reading, *that much* involves word-initial tongue tip raising just prior to [m] that is significantly similar to the tongue tip behavior of *much* in isolation. We argue that English CSD should be analyzed as categorical deletion and that one can only argue for gradient deletion after considering the neutral position of the tongue tip. More generally, this study suggests that arguing for gradience involves complexities beyond merely noting variations in measurements. Therefore, one may conclude categoricity based on Occam's razor and only argue for gradience when alternative explanations have been evaluated and suggested so.

Keywords. coronal stop deletion; categoricity vs. gradience; speech production; comparative technique

1. Introduction. English Coronal Stop Deletion (CSD) — the phenomenon that words like *fact* [fækt~fæk] can be pronounced with or without a coronal stop /t, d/ has been studied in many varieties of English such as American English (Purse 2021), Canadian English (Walker 2012), Singapore English (Lim & Guy 2005), and Southern British English (Baranowski & Turton 2020). One line of research on CSD focuses on its sensitivity to morphological contexts, that deletion is more frequent within monomorphemes, such as *pact*, than across morpheme boundaries, such as *packed* (Guy 1980; Guy & Boyd 1990). There are three accounts for such sensitivity. First, the functional account argues that the past tense morpheme's higher functional load makes it more resistant to deletion (MacKenzie & Tamminga 2021). Second, Baranowski & Turton (2020) believe that morphological structure serves as a constraining factor on the variable deletion rule. Third, Guy (1991) argues that differences at the derivational level for different categories account for the varying deletion rates.

Another line of CSD research is on the question of whether there is categorical or gradient deletion (Scobbie 2007; Purse 2019, 2021). Categorical deletion means that there is always full deletion, and gradient deletion is when there is variation in deletion along a continuum. Among studies that used articulatory data to answer this question, some support categorical deletion, since all speakers pronounced some final /t/s without any tongue tip gesture (Lichtman 2010; Heyward et al. 2014). Others suggest gradience based on the observation that apparent CSD without any tongue tip raising was rarely observed (Purse & Turk 2016; Purse 2019, 2021). Moreover, Browman et al. (1990) argued that there is no deletion rule applied since coronal stops are inaudible when the tongue tip gesture is masked.

A complication with these previous studies is that they overlooked factors other than gradient deletion or masking that can result in an inaudible tongue tip movement. One such factor is

* We are grateful for Dr. Karthik Durvasula's guidance, advice, and support on this project. Also, we would like to thank everyone in the Phono Group at Michigan State University. Many thanks to the audience of the Phonetic Variation session of the 2024 Annual Meeting of the Linguistic Society of America, who provided valuable feedback and inspiration for this work. Authors: Yunting Gu, Michigan State University (yuntinggu@gmail.com) & Ryan Peters, Michigan State University (pete1494@msu.edu).

neutral position, the configuration where articulators are positioned just prior to speaking but different from their location during quiet breathing (Chomsky & Halle 1968). Therefore, the observation of tongue tip raising during an inaudible coronal stop alone may not support gradient deletion since the tongue tip raising may not come from undershoot but from neutral tongue tip position. Another factor is the pause posture, which is when specific articulatory movements occur during pauses at strong prosodic boundaries (Krivokapić et al. 2020). When observing coronal stop deletion that occurs at strong prosodic boundaries, it is important to consider the possibility that inaudible tongue tip raising may come from pause posture rather than gradient deletion. Additionally, the proposed method of comparing tongue tip raising in inaudible coronal stops to a baseline is backed by some observations in articulatory studies (Gelfer et al. 1989; Liu et al. 2022). For example, Gelfer et al. (1989) found that American English speakers produce alveolar consonants with significant lip rounding, and Liu et al. (2022) argued that articulatory studies should use a triplet stimuli set that involves baselines.

The current study addresses this complication of previous studies by examining inaudible tongue tip raising in *that much*. Its tongue tip behavior just prior to the bilabial segment is compared to word initial tongue tip behavior in *much* in isolation. We hypothesize that a) the tongue tip location at maximum constriction will not be significantly different and b) the tongue tip duration — both in terms of tongue tip target and tongue tip gesture — will not be significantly different. If similar tongue tip behavior can be found in both scenarios, then inaudible tongue tip raising in CSD comes from neutral position instead of gradient deletion, which is consistent with the categorical deletion hypothesis. By analyzing articulatory data from the Wisconsin X-ray Microbeam Database (Westbury et al. 1990), we have found that the [m] in *much* has concomitant tongue tip raising when in isolation just as it does following a coronal-final word. This suggests that the tongue tip raising observed in CSD comes from neutral position instead of gradient deletion. We argue, from this, that future studies exploring the gradient vs. categorical deletion question should compare the tongue tip behavior in CSD environments to a baseline. Given that duration comparison of tongue tip involves unsolvable complexity relevant to factors such as speech rate, we resort to Occam’s razor and argue that English CSD should be analyzed as categorical deletion.

2. Methods.

2.1. THE CORPUS AND STIMULI. To evaluate the hypothesis that inaudible tongue tip raising actually comes from the neutral position rather than gradient deletion, we analyzed kinematic data from the Wisconsin X-ray Microbeam Database (Westbury et al. 1990). To collect data for this corpus, microphones were used to record acoustic signals and several pellets were placed on each speaker’s head. Figure 1 shows the positions of the pellets schematically. Three pellets were used as reference points, indicated by Ref — one on the bridge of the nose, the second on buccal surface of the maxillary incisors, and the third either on the nosebridge lower than the first or an arm projecting from a snug-fitting pair of eyeglass frames. To extract information about tongue movement, four pellets, denoted by T1 to T4, were attached along the longitudinal sulcus of each speaker’s tongue. T1 was placed 10 mm posterior to the tongue tip, and T4 was placed about 60 mm posterior to the tongue tip, depending on each speaker’s tolerance. Positions of T2 and T3 were chosen so that the four tongue pellets were equally distanced. To record information of labial articulation, one pellet was attached to the upper lip (UL) and one to the lower lip (LL).

The Wisconsin X-ray Microbeam Database has 118 speech production tasks of various types,

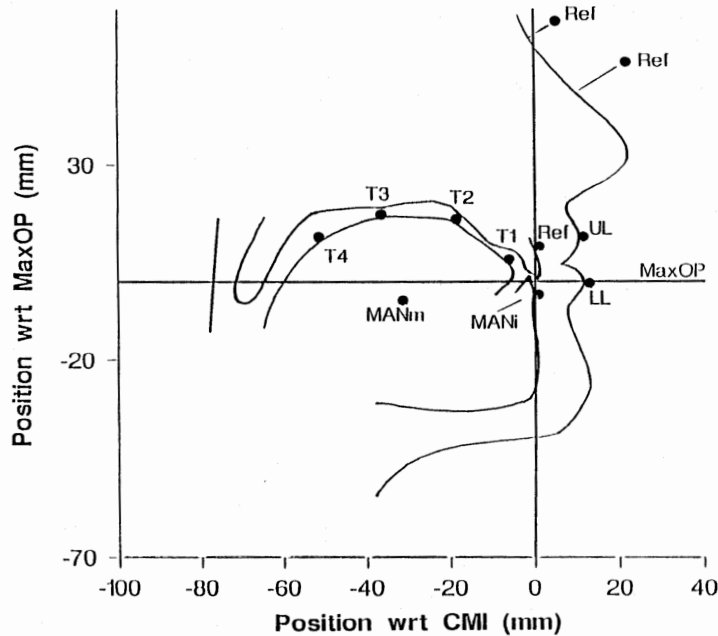


Figure 1. Approximate pellet placement locations. This figure is from Figure 5.2 of the Wisconsin X-ray Microbeam Database manual (Westbury et al. 1990).

including paragraph reading, sentence reading, citation word production, and number sequence production. Speakers represented in the database were recruited from the University of Wisconsin-Madison as well as the surrounding city, and a majority of speakers spoke an Upper Midwest dialect of American English. Altogether, speech production data of 57 different speakers (32 females; 25 males) were included in the database. The median age for the speaker sample was 21.1 years old (female 21.3 years; male 20.8 years). The pair of stimuli for the study consists of *much* in word list reading task 70 and *that much* in sentence reading task 20. There are 3 sentences in task 20, and we analyzed the third sentence, “If I had that much cash, I’d buy the house.”

2.2. DATA ANNOTATION AND ANALYSIS. The data collected by each pellet were visualized by *mview* in Matlab where each pellet’s movement is indicated by one row of curves (Tiede 2005). As in Figure 2, the data were annotated in Matlab using the *lp_fndgest* algorithm of the *mview* package. With this algorithm, the 20 percent thresholds of peak velocity are used for labeling. For each label, the information for the target and gesture is available. In Figure 2, the width of the blue rectangles represents the total duration of the tongue tip gesture, while the solid block in the middle represents the tongue tip target.

In the current study, the coronal stop [t] was annotated by tongue tip, or T1, just prior to [m] in *much*. The [m] in *much* was measured by lip aperture, which was calculated by *mview* (Tiede 2005). Although we intended to annotate data from all 57 speakers, some speakers had missing data for the tongue tip sensor or did not have available data for both tasks. Consequentially, our results are based on the 33 speakers who had data for the target pair.

In terms of analyzing the data, the annotated data were plotted by the *ggplot2* package in R (R Core Team 2017). At the beginning of *much* in both contexts, the tongue tip information at

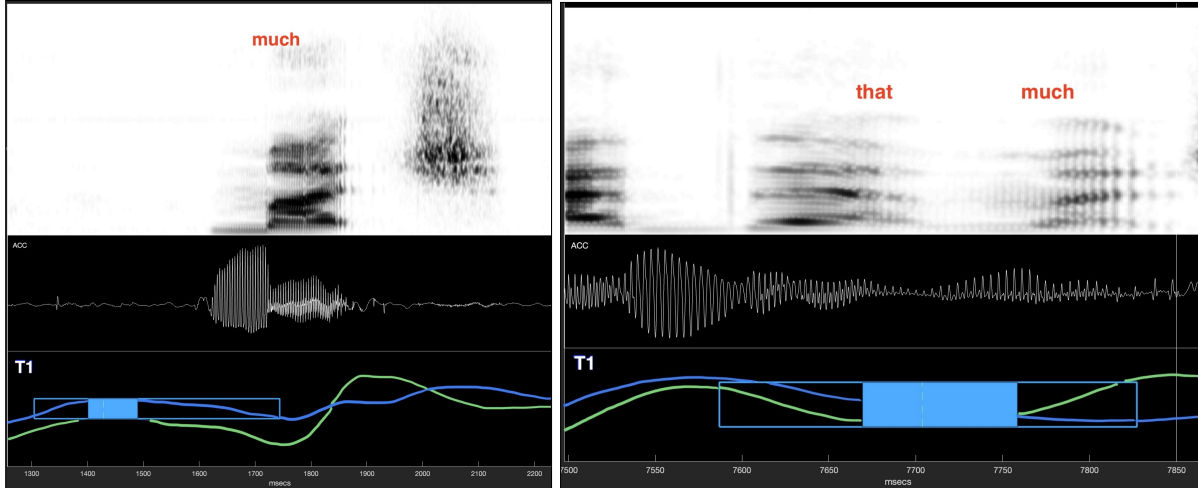


Figure 2. Tongue tip gesture annotation for speaker 39. To the left is *much* in word list reading, and to the right is *that much* in sentence reading. The acoustic information is at the top, the gesture movement information can be found in the panel labeled T1, and the relevant time information is at the very bottom. Tongue tip movement contour is indicated by blue and green curves. The blue curve represents the x-position, and the green curve represents y-position. The width of the blue rectangles represents total duration of the tongue tip gesture, while the solid block in the middle represents the tongue tip target.

vertical and horizontal dimensions was compared respectively using one-sample t-tests in R (R Core Team 2017). The maximum constriction location and duration of the tongue tip gesture in both environments of *much* were compared by visualization and t-tests.

3. Results. The results for comparing tongue tip location, duration, and relativized duration can be found in the following subsections. The results for tongue tip gesture location comparison and tongue tip duration comparison are largely consistent with the categorical deletion hypothesis.

3.1. TONGUE TIP LOCATION COMPARISON. The tongue tip location comparison of *that much* in sentence reading and *much* in isolation for 33 different speakers is shown in Figure 3. The maximum constriction points of the tongue tip gestures for *much* are represented by red triangles, and the corresponding locations of tongue tip gestures for *that much* are shown as black dots. Since there is not any obvious grouping or pattern in terms of the distributions of the tongue tip locations, we conclude that the tongue tip gesture found just prior to the articulation of *much* is similar for both *that much* in sentence reading and *much* in isolation.

To evaluate the statistical significance of this similarity, the vertical and horizontal dimensions of tongue tip maximum constriction points at the beginning of *much* were compared respectively in both contexts. By subtracting the axis information of *much* from that of *that much*, the x and y axis differences were calculated (x-axis difference = $x_{\text{that much}} - x_{\text{much}}$; y-axis difference = $y_{\text{that much}} - y_{\text{much}}$). The differences were then compared to 0 by one-sample t-tests. The results in Table 1 show that both the vertical and horizontal locations of tongue tip just before the [m] in *much* were not significantly different between environments. Also, note that the location difference is about 1 mm (i.e., x-axis: 1.17 mm; y-axis: 0.98 mm) in each dimension. Given that the average distance from the T4 pellet to the tongue apex among the participants of the corpus was

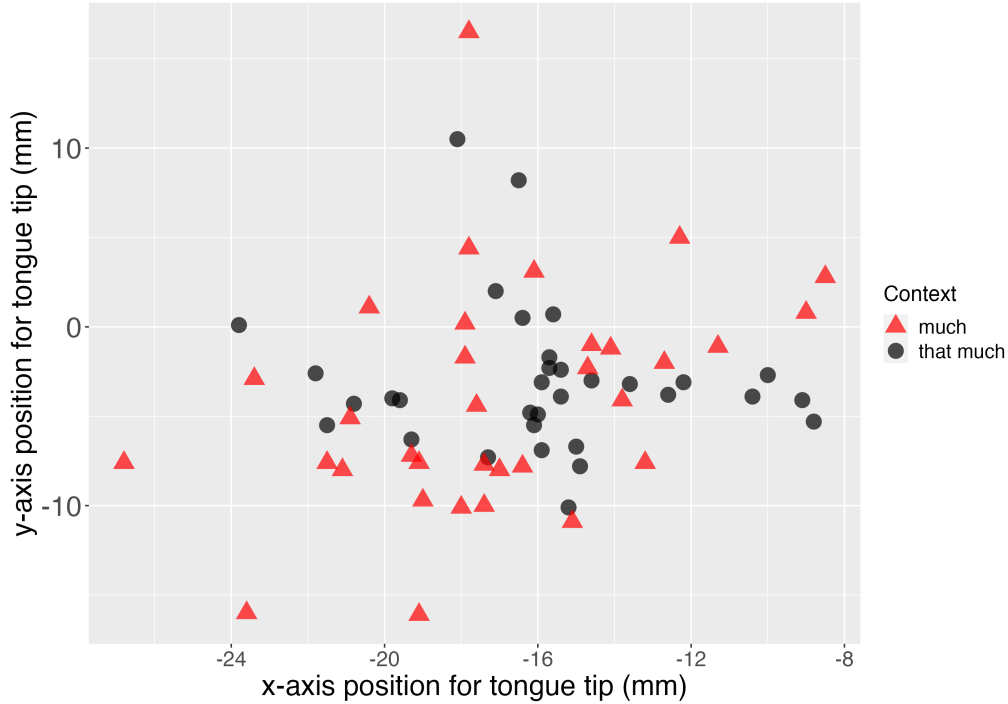


Figure 3. Tongue tip location comparison for all 33 speakers. The maximum constriction points of the tongue tip gestures for *much* are represented by red triangles, and the corresponding locations of tongue tip gestures for *that much* are shown as black dots.

60 mm, the 1 mm difference is considered trivial. Therefore, in terms of location, we conclude that the tongue tip gesture just prior to the articulation of *much* is significantly similar in both CSD and isolated word conditions. In other words, the bilabial gesture in [m] of *that much* has co-occurring tongue tip raising similar to that of *much* in isolation. This further allows us to argue that the tongue tip raising observed at the beginning of *much* in *that much* does not come from gradient but rather categorical deletion.

Data	Estimate	T-Score	DF	P-Value	Confidence Interval (95%)
x-axis difference	1.17 mm	1.56	32	0.13	(-0.35, 2.69)
y-axis difference	0.98 mm	0.91	32	0.37	(-1.22, 3.19)

Table 1. One-sample t-test results for maximum constriction point of tongue tip gesture. DF means Degrees of Freedom. The x and y axis differences were calculated by subtracting the respective position of *much* from that of *that much* — for example, x-axis difference = $X_{that\ much} - X_{much}$.

3.2. DURATIONAL COMPARISON OF TONGUE TIP.

3.2.1. GESTURE DURATION COMPARISON. The tongue tip gesture duration was compared for both environments, and the distribution can be found in Figure 4. We can observe from the figure that the gesture duration for *much*, represented by pink bars, is generally longer than *that much*, represented by blue bars.

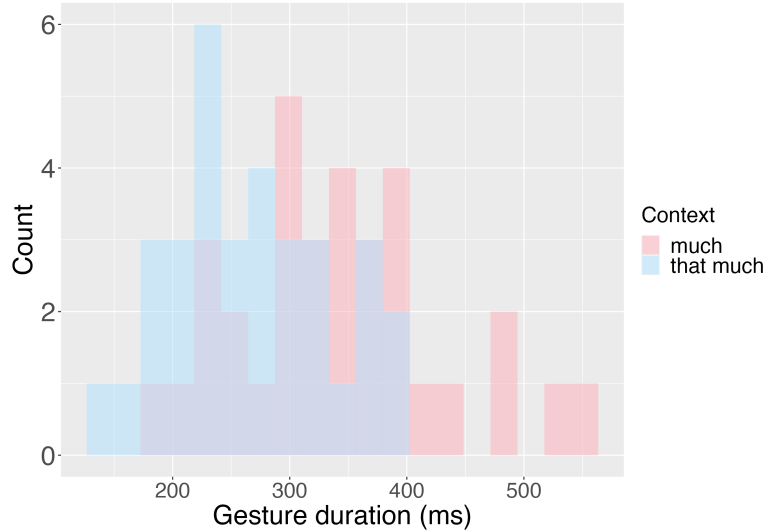


Figure 4. The distribution of tongue tip gesture durations for 33 speakers (binwidth = 23).

The results of the one-sample t-test comparing the durations of the tongue tip gesture just before [m] in *that much* and in isolated *much* are presented in Table 2. These results, including a significant negative estimate, indicate that the gesture has a statistically shorter duration in the CSD condition. We would also like to point out that even though the confidence interval is quite large, the values in the interval are all negative, which suggests that a negative estimate is representative. In short, these results are inconsistent with the gradient deletion hypothesis because one would expect a longer gesture in *that much*, the environment with potential residuals of CSD.

Data	Estimate	T-Score	DF	P-Value	Confidence Interval (95%)
that much - much	-72.81 ms	-3.83	32	0.0006	(-111.58, -34.04)

Table 2. One-sample t-test results for gesture duration difference. DF means Degrees of Freedom.

3.2.2. TARGET DURATION COMPARISON. The distribution of target durations of the tongue tip gesture in both CSD and isolated environments can be found in Figure 5. We can see a large degree of overlap of pink and blue bars, representing target durations for *much* and *that much* respectively.

To evaluate statistical significance, the durational difference of the tongue tip target is compared to 0 using a one-sample t-test. The result in Table 3 shows that even though the target duration of the tongue tip gesture in the CSD environment is about 10 milliseconds longer on average, this difference is not statistically significant. Also, the confidence interval is quite large and spans across negatives and positives, reflecting substantial uncertainty in the estimate. These results show that there is no observed difference between the target duration of tongue tip gesture in *that much* and the target duration of tongue tip gesture co-occurring with [m] of *much* in isolation. Since the tongue tip target durations are similar, observed tongue tip movement in CSD comes from neutral position instead of gradient deletion.

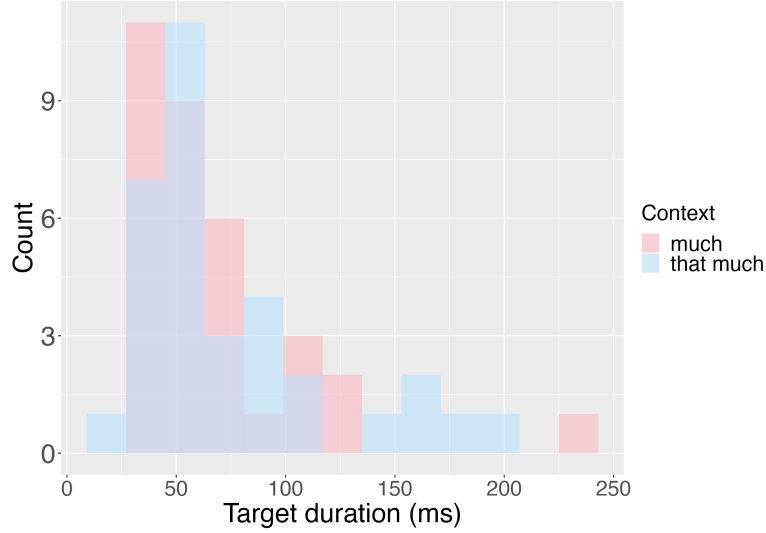


Figure 5. The distribution of tongue tip target duration for 33 speakers (binwidth = 18).

Data	Estimate	T-Score	DF	P-Value	Confidence Interval (95%)
that much - much	9.91 ms	0.91	32	0.37	(-12.28, 32.09)

Table 3. One-sample t-test results for target duration difference. DF means Degrees of Freedom.

3.3. RELATIVIZED DURATION COMPARISON OF TONGUE TIP. Considering domain-initial lengthening, which makes gestures longer (Byrd & Saltzman 2003; Byrd & Krivokapić 2021), it is possible that the tongue tip gesture for *that much* is shorter because it does not have boundary-adjacent lengthening associated with isolated *much*. Moreover, the observed durational difference in tongue tip could be due to speech rate being faster in sentence reading than in word list reading. To compare duration in a way that considers contextual variation, the tongue tip duration is relativized by dividing it by the duration of lip aperture of [m] in *much* as in (1). This relativization has been applied to both tongue tip gesture and tongue tip target.

$$(1) \text{ Relativized tongue tip duration} = \frac{\text{Tongue tip duration}}{\text{Lip aperture duration of [m] in } \textit{much}}$$

Note that dividing the tongue tip gesture duration by lip aperture duration of [m] can only be an effective normalization method if all kinds of gesture — in this case, lip aperture and tongue tip — change their duration to the same degree with respect to speech rate variation. Assuming the premise that all gestures do react similarly to speech rate variation, we subtracted the relativized tongue tip duration of *much* from the relativized tongue tip duration of *that much* as in (2) and compared the difference to 0 using a one-sample t-test.

$$(2) \frac{\text{Tongue tip duration of } \textit{that much}}{\text{Lip aperture duration of [m] in } \textit{that much}} - \frac{\text{Tongue tip duration of } \textit{much}}{\text{Lip aperture duration of [m] in } \textit{much}}$$

The plots for the relativized duration of tongue tip gesture and tongue tip target can be found in Figure 6 where pink indicates the information for *much*, and blue shows the relativized duration for *that much*. The histogram in Figure 6a shows that while there is overlap in the duration of both contexts, tongue tip gesture is relatively longer for *that much*. Similarly, Figure 6b shows more clearly that the target of the tongue tip gesture for *that much* is longer than that in *much*.

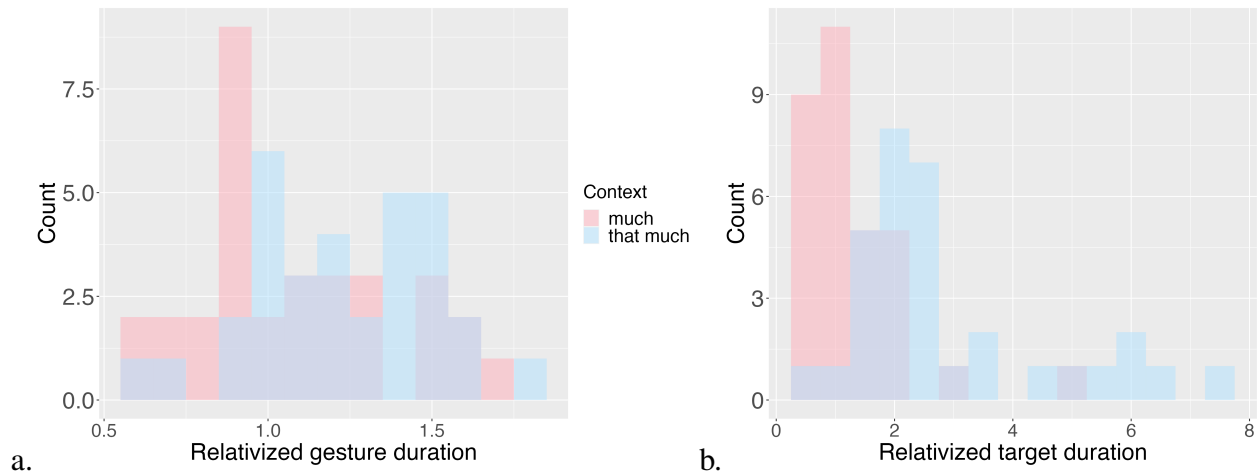


Figure 6. The distribution of relativized duration. Pink represents the relativized duration for *much*, and blue indicates the relativized duration for *that much*. Plot a: the distribution of relativized tongue tip gesture duration for 32 speakers (binwidth = 0.1). Plot b: the distribution of relativized tongue tip target duration for 32 speakers (binwidth = 0.5).

The one-sample t-test results for the relativized gesture and target duration of tongue tip can be found in Table 4. We can see that for both gesture and target comparison, the relativized duration for *that much* is relatively longer than the relativized duration for *much*. While both comparisons show statistical significance, target duration exhibits a more significant difference.

Comparison	Data	Estimate	T-Score	DF	P-Value	CI (95%)
Gesture	that much - much	0.15	2.23	31	0.03	(0.01, 0.29)
Target	that much - much	1.64	4.88	31	0.00003	(0.95, 2.32)

Table 4. One-sample t-test results for relativized duration difference. DF means Degrees of Freedom. CI means confidence interval.

One may interpret the observation of a longer normalized duration of tongue tip gesture in *that much* as gradient coronal stop deletion; however, the justification behind this interpretation remains unclear. Without an independent study to evaluate whether or not different gestures react to speech rate similarly, the observation is inconsistent with the gradient deletion hypothesis.

3.4. SUMMARY OF RESULTS. We found that the maximum constriction point of tongue tip gesture, as well as the target duration of tongue tip movement, were not significantly different in *much* and *that much*. We also found that the gesture duration comparison result is inconsistent with the gradient deletion hypothesis. Moreover, even though the relativized duration comparison results seem to support the gradient deletion hypothesis, the premise for the argument lacks proper support.

Conclusion and discussion. There is evidence from the current study, of both the location and duration aspects, to support categorical deletion of English coronal stops. Since our results are either inconsistent with the gradient hypothesis or involve difficulty in interpretation, we resort to Occam's razor and argue that English coronal stop deletion is categorical, not gradient. This study strengthens the argument that the neutral position serves as a baseline of targeted articulatory movement, and it suggests that neutral position should be considered in future articulatory studies.

Admittedly, there are limitations in our study. In our stimuli, we only considered one word pair, and this word pair was not the prototypical CSD consonant cluster. We also would like to caution the reader about overestimating the statistical significance of the study. Even though we analyzed articulatory data from 33 speakers, there were cases in which the confidence interval was large. Having more repetitions from each speaker may narrow the confidence interval. Future research is necessary to test CSD on more word pairs with more repetition, as well as on more varieties of English, using an experiment paradigm that considers factors such as neutral position.

Despite the limitations, our study demonstrates how additional factors besides gradient deletion must be considered when using articulatory data to probe gradience vs. categoricity. When presenting our finding of word-initial bilabial segments having concomitant tongue tip raising, our crucial claim is more about the merit of a comparative paradigm than about emphasizing the neutral position. It is logically possible that in our study, there were cases where the gesture detected was not a gesture but rather noise. Therefore, future evaluation is necessary to distinguish between the different possible sources of the observed tongue tip raising.

In general, the current study uses the comparative paradigm to address the question of whether there is categorical or gradient coronal stop deletion in English. It shows the complexity of interpreting articulatory observation, and one should only argue for gradience when alternative explanations have been carefully evaluated and suggested so.

References

- Baranowski, Maciej & Danielle Turton. 2020. TD-deletion in British English: New evidence for the long-lost morphological effect. *Language Variation and Change* 32(1). 1–23. <https://doi.org/10.1017/S0954394520000034>.
- Browman, Catherine P, Louis Goldstein et al. 1990. Tiers in articulatory phonology, with some implications for casual speech. In John Kingston & Mary E. Beckman (eds.), *Papers in laboratory phonology I: Between the grammar and physics of speech*, 341–376. Cambridge: Cambridge University Press.
- Byrd, Dani & Jelena Krivokapić. 2021. Cracking prosody in articulatory phonology. *Annual Review of Linguistics* 7. 31–53. <https://ssrn.com/abstract=3778503>.
- Byrd, Dani & Elliot Saltzman. 2003. The elastic phrase: Modeling the dynamics of boundary-adjacent lengthening. *Journal of Phonetics* 31(2). 149–180. [https://doi.org/10.1016/S0095-4470\(02\)00085-2](https://doi.org/10.1016/S0095-4470(02)00085-2).
- Chomsky, Noam & Morris Halle. 1968. *The sound pattern of English*. New York: Harper & Row.
- Gelfer, Carole E., Fredericka Bell-Berti & Katherine S. Harris. 1989. Determining the extent of coarticulation: Effects of experimental design. *The Journal of the Acoustical Society of America* 86(6). 2443–2445. <https://doi.org/10.1121/1.398452>.
- Guy, Gregory R. 1980. Variation in the group and the individual: The case of final stop deletion. In William Labov (ed.), *Locating language in time and space*, 1–36. New York: Academic Press.

- Guy, Gregory R. 1991. Contextual conditioning in variable lexical phonology. *Language Variation and Change* 3(2). 223–239.
- Guy, Gregory R. & Sally Boyd. 1990. The development of a morphological class. *Language Variation and Change* 2(1). 1–18.
- Heyward, J., A. Turk & C. Geng. 2014. Does /t/ produced as [a] involve tongue tip raising? articulatory evidence for the nature of phonological representations. Presented at the 14th Conference on Laboratory Phonology, Tokyo.
- Krivokapić, Jelena, Will Styler & Benjamin Parrell. 2020. Pause postures: The relationship between articulation and cognitive processes during pauses. *Journal of Phonetics* 79. 100953. <https://doi.org/10.1016/j.wocn.2019.100953>.
- Lichtman, Karen. 2010. Testing articulatory phonology: Variation in gestures for coda /t/. Presented at the Illinois Language and Linguistic Society Conference, University of Illinois.
- Lim, Lauren T. & Gregory R. Guy. 2005. The limits of linguistic community: Speech styles and variable constraint effects. *University of Pennsylvania Working Papers in Linguistics* 10(2). 13.
- Liu, Zirui, Yi Xu & Feng-fan Hsieh. 2022. Coarticulation as synchronised CV co-onset – Parallel evidence from articulation and acoustics. *Journal of Phonetics* 90. 101116. <https://doi.org/10.1016/j.wocn.2021.101116>.
- MacKenzie, Laurel & Meredith Tamminga. 2021. New and old puzzles in the morphological conditioning of coronal stop deletion. *Language Variation and Change* 33(2). 217–244.
- Purse, Ruaridh. 2019. The articulatory reality of coronal stop deletion. *International Congress of Phonetic Sciences (ICPhS)* 19. 1595–1599.
- Purse, Ruaridh & Alice Turk. 2016. /t, d/ deletion: Articulatory gradience in variable phonology. Presentation at The 15th Conference on Laboratory Phonology (LabPhon15, Ithaca, NY).
- Purse, Ruaridh. 2021. *The role of morphological structure in phonetic variation*. Philadelphia: University of Pennsylvania dissertation.
- R Core Team. 2017. *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. <https://www.R-project.org/>.
- Scobbie, James M. 2007. Interface and overlap in phonetics and phonology. In Gillian Ramchand & Charles Reiss (eds.), *The Oxford handbook of linguistic interfaces*, 17–52. Oxford: Oxford University Press.
- Tiede, Mark. 2005. Mview: Software for visualization and analysis of concurrently recorded movement data. New Haven, CT: Haskins Laboratories.
- Walker, James A. 2012. Form, function, and frequency in phonological variation. *Language Variation and Change* 24(3). 397–415.
- Westbury, John, Paul Milenkovic, Gary Weismer & Raymond Kent. 1990. X-ray microbeam speech production database. *The Journal of the Acoustical Society of America* 88. S56. <https://doi.org/10.1121/1.2029064>.