Dissertation Proposal:
Categoricity and Gradience in Coronal Stop Deletion

Ruaridh Purse

Supervisor: Meredith Tamminga
Committee: Mark Liberman (chair), Jianjing Kuang, David Embick

Abstract

Coronal Stop Deletion (CSD) in English has long enjoyed pride of place as the test case for approaches to variable phonology. However, CSD’s realisation as a categorical alternation has always been assumed, and not demonstrated. This dissertation seeks to provide a unified analysis of CSD, exploring the details of its production and perception for clues about its representation.

In this proposal, I present pilot EMA data from five speakers, in which it seems that complete deletion with no tongue tip raising is rare but does exist. Moreover, some individuals exhibit articulatory profiles that could be consistent with a structure non-preserving variable allophony analysis. Simultaneously, effects of morphology, task, and speech rate on a gradient measure of tongue tip raising suggest current phonological analyses of CSD still do not capture the whole story.

The core of the dissertation will involve gathering a large-scale Electromagnetic Articulography (EMA) corpus of semi-spontaneous speech. This data will be crucial in addressing questions of categoricity and gradience in the implementation of CSD, as well as individual differences along these parameters. The articulatory core of the dissertation will additionally be supplemented with a perception study and a laboratory acoustic production study. The perception study should confirm whether listeners can tell the difference between coronal stops that would traditionally be classified as deleted but differ in terms of tongue tip raising. The acoustic production study will investigate an alternative representation for CSD that allows for some level of gradience in tongue movements without challenging the modularity of the grammar; namely that words may differ in their prosodic structure, giving rise to different rates of lingual undershoot.

1. Introduction

English Coronal Stop Deletion (CSD, sometimes called ‘/t,d/ Deletion’) is the variable surface absence of an underlying word-final coronal stop following a consonant. Since its first description as cluster simplification from Labov et al. (1968), it has perhaps become the most thoroughly investigated phenomenon in English variationist sociolinguistics. As it has typically been coded in sociolinguistic research, CSD could be formulated as a phonological rule such as in (1).

\[ C^{\text{COR}} \rightarrow \emptyset /C \# \]  

However, both the representation in (1) and the name (‘Deletion’) contain assumptions about CSD for which there is a dearth of evidence. Specifically, the notion that a variable /t/ is probabilistically realised as discrete categories [t] or [Ø] has the potential to be a serious oversimplification of the details of phonetic implementation. In this document I present preliminary results from using Electromagnetic Articulography to explore the articulation. I outline plans to scale up this methodology to many more speakers, and to supplement it with perception and acoustic production studies. This combination of approaches will be suitable for addressing longstanding questions about the representation of CSD.
1.1. ‘Coronal Stop Deletion’

In the last 50 years CSD has been recognised as something of a pan-English phenomenon, with accounts of its presence in varieties of English around the world (Labov, 1972; Wolfram, 1969; Fasold, 1972; Guy, 1980; Santa Ana, 1991; Patrick, 1991; Holmes and Bell, 1994; Tagliamonte and Temple, 2005; Lim and Guy, 2005; Gut, 2007; Guy et al., 2008; Hazen, 2011; Walker, 2012; Hansen Edwards, 2016). Among this plethora of analyses of CSD conducted across Englishes, there are some striking consistencies. In almost every study of the phenomenon, the rate of deletion has been found to strongly depend on phonological and morphological context. One basic effect of phonological context is that CSD is far more common when an eligible coronal stop is followed by a consonant (e.g. west square) than a vowel (e.g. west avenue). Several accounts attribute this effect to the variable resyllabification of a coronal stop to be the onset to a following vowel (Guy, 1991; Kiparsky, 1993; Reynolds, 1994). This resyllabification will necessarily bleed CSD, which only targets coda stops. The following segment effect has additionally been found to be systematically modulated by the strength of the intervening syntactic boundary (Tamminga, 2018). Tamminga argues that this is the result of an extra-grammatical effect of production planning. In other words, speakers are less likely to plan the realisation of a coronal stop in conjunction with subsequent segments when there is a strong syntactic boundary that intervenes, compared to when the intervening boundary is weaker. This interpretation must be understood in categorical (i.e. phonological) terms; segments are or are not planned in conjunction. This is in contrast with a hypothetical coarticulatory effect that might be stronger or weaker depending on how quickly segments are articulated in sequence.

The simple observed effect of morphology on CSD is that deletion occurs more frequently in monomorphemes (e.g. pact) than in words where it constitutes an -ed suffix to mark a passive or preterite form (e.g. packed). Some early attempts to explain this phenomenon have a ‘functional’ flavour such as that invoked by Kiparsky’s (1972) ‘distinctness conditions’: the pressure to retain informative structure – like a suffix that marks the past tense – offsets the pressure to economise on articulatory effort. Others are more strongly informed by morphophonological theory and attribute differences in deletion rates to structural aspects of word-formation. This is true of Guy’s (1991) Exponential Hypothesis, wherein monomorphemes undergo CSD at a higher rate because it is repeatedly applied at multiple levels of a stratified morphophonology (e.g. Kiparsky, 1982), whereas an -ed suffix is inserted late and only eligible for one level’s CSD process. For the many theorists who assume strict feed-forward modularity in speech production (e.g Bermúdez-Otero, 2007), this morphological conditioning is strong evidence for the phonological character of CSD. Under a view where linguistic ‘modules’ (such as those that generate outputs for syntax and phonology) are domain specific, mandatory, and strictly ordered (Fodor, 1983), morphology and phonetics do not share an interface: phonology must intervene. As a result of this and the facts about phonological conditioning, CSD has had a special position as a primary case study for many models of variable phonology.

An assumption that is frequently made, to varying degrees of explicitness, is that phonological and phonetic operations and objects differ crucially with regard to categoricity1 and gradience (Scobie, 2007). While matters of phonetic implementation are typically characterised in terms of behaviour along various continuous dimensions in acoustic or articulatory space, phonology is represented in terms of computations over strings of discrete symbols comprised of sets of discrete features (cf. work on gradience in phonological representation and derivation, e.g. Hayes, 2000; Smolensky and Goldrick, 2016; Rosen, 2016; Stevenson and Zammuner, 2017). A similar assumption underlies discussions of the character of CSD. When the inability to audibly perceive an underlying coronal stop is attributed to a phonological process, this is a qualitatively different belief than the idea that gradient phonetic lenition or miscellaneous coarticulatory processes – the result of moving articulators to execute a phonological plan in real time – have rendered the stop inaudible. Temple (2014) finds reason to favour the latter

---

1I use ‘categorical’ to refer to representations and implementations that comprise a set of discrete categories, rather than falling across a continuum. This terminology follows a growing body of work on issues of categoricity and gradience in variable phenomena. ‘Categorical’ will never be used to mean ‘invariant’.
perspective in the multitude of confounds that exist for researchers basing judgments on acoustic data, as has traditionally been the norm. Specifically, Temple identifies several environments that are susceptible to neutralising or articulatory overlap phenomena in her data from York English, such that for many cases it is no longer clear whether they can be treated as a normal candidate for CSD. Further, Temple claims that the phonetic detail of CSD tokens is just like that of any other ‘connected speech process’, and while this does not preclude structured implementation it entails that coronal stops should not receive a special abstract phonological analysis. However, Temple does not provide articulatory evidence for this claim.

If CSD is recast as a connected speech process resulting from gradient phonetic effects that are only perceived as categorical, the conditioning of morphology on deletion rates is particularly difficult to explain. If morphological information is relevant for phonetic implementation, we might predict more widespread observation of this phenomenon (e.g. VOT in nouns as conditioned by grammatical gender, speech rate as conditioned by tense). One potential solution is that the effect of morphological class might be an artifact of another, correlated, factor. Bybee (2002) suggests that this factor is lexical frequency. It has been observed that lenition phenomena occur more frequently and to a greater extent in words that are more frequent than in words that are less frequent (Pierrehumbert, 2002). This could be interpreted as a kind of ‘functionalist’ phenomenon since frequent words are more predictable on the part of the listener while infrequent words may require more articulatory clarity to disambiguate them from the more frequent words that are anticipated by default. Bybee (2002) claims that monomorphemes are more frequent on average than passive or preterite forms ending in -ed, and that monomorphemes more frequently appear in environments where the coronal stop is particularly susceptible to lenition. Therefore, monomorphemic words accumulate a representation in which the absence of an audible stop is in keeping with a large number of the instances in which the word was heard or produced, while preterite and passive forms do not.

1.2. Articulatory sociophonetics

Articulation and acoustics are associated non-linearly: this much is well understood. Stevens (1972, 1989) describes the relation between articulation and acoustics as ‘quantal’ such that some large articulatory movements have little acoustic consequence, while some small adjustments are used to produce qualitative (even phonemic) differences. The case of stop consonants is a particularly strong example of this. Only a complete closure at the alveolar ridge will produce the characteristic stop and release burst of a [t] or [d]. However, there is a wide range of tongue tip positions along a superior-inferior dimension, each of which (as long as no closure is made) produces an acoustic signal that is no more or less stop-like than the next. Consequently, a coronal gesture towards a stop closure that is undershot by even a very small distance along this axis could be acoustically indistinguishable from a case of true categorical deletion with no residual articulation of a coronal stop. Thus, controversy surrounding issues of categoricity and gradience – phonology and phonetics – in CSD necessitates the exploration of articulatory data. This will allow us to observe the range of articulations underlying the categorical perception of stop deletion or retention that are inextricable from the acoustic signal.

The research that has been conducted on this narrow topic has yielded some mixed results. In a classic work from Browman and Goldstein (1990), X-Ray Microbeam data revealed evidence of inaudible tongue tip raising for word-final coronal stops. For the two instances of the sequence perfect memory where this was observed, Browman and Goldstein (1990) conclude that the stop was not rendered inaudible from a deletion rule, but rather the difficult coordination of articulators across time. Either the coronal closure and release were either masked by the temporal overlap of adjacent labial and dorsal closures, or the speaker undershot their target of the alveolar ridge and never made a complete closure in the first place. Similarly, Purse and Turk (2016) observe that apparent CSD without tongue tip raising is rare in the ESPF DoubleTalk Corpus. However, for 9 tokens (25% of all inaudible coronal stops) speakers appeared to produce no tongue tip raising, and some of these featured distinct downward tongue tip movement. It is still not clear, however, whether these 9 tokens constitute categorical deletion or just one extreme of a continuum of lenition. On the other hand,
Lichtman (2010) conducted an investigation of the Wisconsin Microbeam Database and a follow-up EMA study, and reports widespread articulatory categoricity in coda /t/ deletion such that all speakers exhibit some tokens with no residual linguoalveolar gesture. It is important to note, however, that the majority of tokens included in this study are singleton stops, not part of consonant clusters as is typically held to be the environment that is eligible for CSD. As such, some of these instances may be the result of different processes such as /t/-glottaling. Indeed, Heyward et al. (2014) provide evidence that instances of /t/-glottaling in the ESPF DoubleTalk Corpus typically feature no evidence of residual tongue tip raising. Instead, this phenomenon resembles true categorical allophony in that /t/ loses its anterior place of articulation and is realised on another tier entirely.

The paucity of research making use of articulatory data is a problem for sociophonetics in general, which lags behind the fields of traditional phonetics and speech pathology in this respect. Indeed, in Thomas’s (2008) review of instrumental phonetic techniques for sociolinguistics, there is no mention of methodologies for measuring articulation. One major obstacle to the incorporation of these methodologies in sociolinguistics is the normalisation of measures across multiple speakers with different anatomies. Studies in articulatory sociophonetics must explore creative solutions for making meaningful comparisons between observations from various individuals. Another potential obstacle lies in the elicitation of the naturalistic speech that is of primary interest to sociolinguists. Where Labov’s (1972) style-shifting paradigm classifies all laboratory speech as a context in which attention-paid-to-speech is bound to be high, any setup for physically observing articulators is likely only to exacerbate this problem. However, Boyd et al. (2015) report that speech from laboratory tasks is largely comparable with speech produced in a sociolinguistic interview. While they do not include a comparison with a context in which articulatory data is gathered, this finding is encouraging in its suggestion that the effects of researcher observation are fairly consistent. It seems likely that even in the present study the effects of conspicuous articulatory methodology should not be so egregious as to level out all the potential variation.

Some areas of sociophonetics have already benefitted from articulatory studies to explore the reality of speakers’ production strategies. For example, it is understood that an English approximant /r/ can be produced with various covertly allophonous tongue shapes, which can be broadly categorised in terms of a bunched versus retroflex taxonomy (Delattre and Freeman, 1968), but which are perceptually indistinguishable (Twist et al., 2007). Using Ultrasound Tongue Imaging (UTI), it has been demonstrated that speakers adapt their /r/ articulation to produce the least effortful allophone given the phonetic context (Stavness et al., 2012) and perturbations of their articulators (Tiede et al., 2011), and that children explore different articulations during acquisition (Magloughlin, 2016). Further, /r/ articulation has been shown to be class stratified in Scottish English (Lawson et al., 2011), and to play a key role in the actuation of s-retraction in North American English (Mielke et al., 2010; Baker et al., 2011; Mielke et al., 2016).

As well as revealing covert ranges of articulatory movement (e.g. incomplete tongue tip raising for coronal stops), and covert allophony (e.g. different lingual configurations for English approximant /r/), articulatory data is particularly informative regarding the matter of timing in speech production. Timing in speech is relative, such that we can conceive of it as the covariation of multiple objects in time, or as variation along a continuum of time between gestures. Relative timing of lingual gestures has been revealed to be a key locus of variation for /l/ darkening (Sproat and Fujimura, 1993; Bermúdez-Otero and Trousdale, 2012; Turton, 2017) and Scottish English /r/ pharyngealisation (Lawson et al., 2018). In these cases, a delayed anterior gesture – sometimes until after voicing ends – leads to a darker /l/ and a more pharyngealised /r/ respectively.

Timing of gestures is also the main locus of variation available in an Articulatory Phonology framework (Brownman and Goldstein, 1990). Here, temporal overlap of gestures associated with adjacent speech sounds gives rise to coarticulatory outcomes or the masking or omission of a gesture. Further, Davidson and Stone (2004) explore a timing analysis for English speakers’ production of excrecent vocoids that variably appear in phonotactically illegal clusters. They conclude that this is indeed a case of gestural mistiming based the absence of evident vowel targets in their UTI data. This result implies, as Brownman and Goldstein (1990) predict, a continuum of overlap between any two given
gestures. At one end of a continuum of this kind of overlap, where gestures are maximally separated in time, there is the potential for apparent vowel epenthesis phenomena as in Davidson and Stone (2004). The other, extreme, end of this kind of continuum of gestural overlap is presumably metathesis, in which the underlying order of speech sounds is reversed on the surface. In cases of both variable vowel epenthesis and elision, and variable metathesis, a clear research goal should be to observe how variants are distributed across a continuum of gestural overlap. A unimodal distribution in the degree of gestural overlap would suggest that the variable perception of vowels or reordering of segments is a product of a gradient timing relationship between gestures. On the other hand, a bimodal distribution of these results would suggest the existence of discrete categories. In the latter case we could surmise that speakers have separate targets for multiple potential surface forms. Articulatory data is crucially poised to provide this kind of evidence. With it, a researcher can disentangle simultaneously produced gestures and is not limited to data on these gestures’ completion but also their onset or ‘Maximum Acceleration Event’ (Perkell and Matthies, 1992) as may be most informative concerning the planning and execution of articulatory timing.

In this proposal, I present a 5-speaker pilot study using EMA to explore the production of underlying coronal stops that are potential CSD targets. The central component of the completed dissertation will be the dataset yielded from applying this methodology to a much larger number of speakers. This will allow us to seek answers to questions concerning the representation of CSD that are raised in the pilot but for which the data are inconclusive. This EMA corpus will be supplemented with two complementary studies. One will explore listeners’ capacity to perceive elements of articulatory detail that are identified in the EMA corpus. The other will explore the potential for an alternative framing of well-attested patterns in CSD that may better capture the articulatory facts: namely, that speakers’ rate of apparent CSD could be a correlate of different prosodic structures associated with different word types.

2. Pilot Study Methods

2.1. Procedure

Synchronised acoustic and articulatory recordings were collected using an NDI Wave Electromagnetic Articulograph and a microphone, through NDI’s native software NDI Wavefront. Electromagnetic Articulography (EMA) sensor coils were adhered to key oral articulator points at the tongue tip (TT), tongue dorsum (TD), and lower lip (LL), as well as a reference point on the upper incisors (UI) using a non-toxic high viscosity cyanoacrylate oral adhesive. Three further reference sensors were aligned to each participant’s left and right mastoids and bridge of nose using a lensless spectacle frame that was held in place with surgical tape. The spectacle-mounted sensors were used to define a participant’s sella-nasion plane at each frame of time, and new axes (inferior-superior, posterior-anterior, left-right) were created according to this plane to correct for participant head movement. The origin of each new axis was then aligned to the UI sensor to improve interpretability.

All 5 participants are native speakers of North American Englishes. Each one performed several tasks designed to elicit naturalistic speech. These were a Map Task (participants describe a route on a map so that an interlocutor can draw it), a Semantic Differential Task (participants explain the difference between near-synonyms), two Reading Passages, and finally a Wordlist. Tasks were consistently ordered in the way presented here under the assumption that this creates a continuum of style such that each task evokes a higher degree of metalinguistic awareness than the last. This follows some classic sociolinguistic methodology (Labov, 1972) in which researchers attempt to tightly control and gradually increase speakers’ degree of self-monitoring across the duration of the experiment. This is an important consideration given the observation that articulatory methodologies are likely to already induce a high level of self-monitoring and metalinguistic awareness. Stimuli for all tasks were designed so as to require participants to produce as many critical items as possible, where a critical item is a
word with an underlying word-final stop following another consonant, with no other adjacent coronal segments.

2.2. Data Manipulation

As previously mentioned, some by-speaker and by-token normalisation is required in order to compare observations from different speakers. Sensor positions cannot be precisely equivalent because the size and shape of each speaker’s body is not the same. For each speaker, a measure of the greatest TT height (mm from plane at UI) for a coronal stop closure was recorded and represented a speaker-specific maximum (MAX). For every token, the TT tangential velocity minima immediately preceding and following (A and B, respectively) are defined as coordinates in Time (s) and TT height (mm), and used to define the baseline AB. The TT tangential velocity minimum corresponding to coronal stop articulatory target for this token (T), where the tongue has been raised, is also recorded in terms of TT height and Time. These values are then used to calculate the distance from AB to T (h), which is then redefined as a proportion of the distance from AB to that speaker’s MAX (H). This normalised measure of raising is always \( \leq 1 \), since it is calculated from \( h/H \). Figure 1 is a schematic with the component parts for this normalisation procedure, for a TT height trajectory across time corresponding to a hypothetical coronal stop closure and release.

![Figure 1: Schematic for calculating normalised measure of tongue tip raising.](image)

The normalised measure that is described here could be thought of as a measure of degree of effort expended to reach a speaker-specific maximum TT height. This is exceptionally relevant to CSD as evaluated as a lenition phenomenon, especially when lenition is narrowly construed in terms of a reduction in the magnitude of potential articulatory movement. Values of this measure are no longer particularly informative about absolute tongue tip height, because they are strongly affected by the height of the baseline AB from which TT raising takes place. In other words, we may observe very little raising to reach a relatively high peak if the corresponding baseline is already high. This means that measures of both raising – a normalised measure of the proportion of maximum articulatory movement from a baseline – and raw height, are crucial and crucially different for the present study\(^2\).

\(^2\)Therefore, ‘raising’ and ‘height’ are not used interchangeably in this paper.
3. Pilot Study Results

3.1. Is ‘Coronal Stop Deletion’ implemented categorically?

3.1.1. Auditory coding

This study is primarily an investigation of the articulatory reality of word-final coronal stops. However, in order to speak to the previous literature on CSD, the data must first be evaluated in these traditional terms. Table 1 shows the rates of CSD in each of the basic morphological classes this paper considers (monomorphemes, ‘complex’ regular passive and preterite forms, and semiweak past forms). These tokens were coded according to auditory and spectrographic cues to the presence or absence of a coronal stop. In terms of phonetic environment, only tokens that were immediately followed by another coronal stop or an interdental fricative, since these are almost always neutralising (Temple, 2009).

<table>
<thead>
<tr>
<th></th>
<th>Retained</th>
<th>Deleted</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mono</td>
<td>316 (52%)</td>
<td>286 (48%)</td>
<td>602</td>
</tr>
<tr>
<td>Complex</td>
<td>297 (74%)</td>
<td>101 (26%)</td>
<td>398</td>
</tr>
<tr>
<td>Semi</td>
<td>24 (72%)</td>
<td>9 (28%)</td>
<td>33</td>
</tr>
<tr>
<td>Total</td>
<td>637 (63%)</td>
<td>396 (37%)</td>
<td>1033</td>
</tr>
</tbody>
</table>

Table 1: Auditory/acoustic coding of coronal stop retention and deletion.

The well-attested effect of morphological class on CSD is also found in this data, such that CSD ostensibly occurs at a significantly higher rate in monomorphemes than Complex words ($\chi^2 = 49.5, p < 0.00001$). Indeed, according to these results, CSD occurs almost twice as frequently in monomorphemes as in passive or preterite forms. Semiweak past forms appear to pattern with Complex forms, but the sample size for this subset is too small to be particularly informative.

3.1.2. Articulatory ‘zeroes’

The bulk of the analysis in this paper concerns tokens of underlying coronal stops with no adjacent coronal segments. The rate of perceived CSD in this subset, based on auditory and spectrographic cues, is a respectable 24%. However, it is not clear from this evidence whether any instance of apparent CSD actually constitutes an absence of any attempt to produce a stop, as the classic analysis implies. In the present analysis, there were 15 tokens in which there was no evidence of tongue tip raising. In each of these, the TT trajectory during interval of time corresponding to an underlying coronal stop stayed level with AB or moved downwards. None of these featured audible stops and would have been judged to be ‘deleted’ in a traditional CSD analysis. Figure 2 is one such token.

Table 2 shows the rates of coronal stops in each grammatical class that were audible or inaudible, and within the inaudible class how many did or did not exhibit TT raising (raising ≤ 0).

<table>
<thead>
<tr>
<th></th>
<th>Audible</th>
<th>Inaudible</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>+ Raising</td>
<td>− Raising</td>
<td></td>
</tr>
<tr>
<td>Mono</td>
<td>123 (79%)</td>
<td>29 (19%)</td>
<td>4 (3%)</td>
</tr>
<tr>
<td>Complex</td>
<td>139 (72%)</td>
<td>42 (22%)</td>
<td>11 (6%)</td>
</tr>
<tr>
<td>Semi</td>
<td>13 (93%)</td>
<td>1 (7%)</td>
<td>0 (0%)</td>
</tr>
<tr>
<td>Total</td>
<td>275 (76%)</td>
<td>72 (20%)</td>
<td>15 (4%)</td>
</tr>
</tbody>
</table>

Table 2: Articulatory categorisation of perceived coronal stop retention and deletion.

Out of 87 tokens that were inaudible, just 15 (17%) appear to be true ‘articulatory zeroes’ with no evidence of TT raising to create a coronal closure. If similar articulatory profiles belie traditional
acoustic research on CSD, it may mean that the rates of deletion taken as the output of phonological processes are vastly overestimated. Generalising these results by morphological class, we might predict that only 12% of monomorphemes and 26% of -ed suffixed forms where deletion has been observed can actually be described as such. On the other hand, this rate of ‘articulatory zeroes’ far outstrips anything that has been observed before (Browman and Goldstein, 1990; Purse and Turk, 2016).

There is no apparent asymmetry in the distribution of articulatory zeroes amongst morphological classes. Such a thing might be expected if we were to attribute the absence of audible coronal stops to different processes with specific inputs. Such a pattern may become evident given a larger sample size, but a sufficient sample will be a challenge to obtain given the nature of articulatory data and the apparent rarity of articulatory zeroes as a phenomenon. It should also be noted that, unlike the analysis that included all non-neutralising phonetic environments in §3.1.1, the analysis that is limited to environments with no adjacent coronal segment does not show the normal morphological conditioning on rate of inaudible stops. That is, there is not a greater proportion of inaudible stops in monomorphemes compared to complex words. This, too, may be a function of the relatively small sample size. It is not cause for too much concern, given that the expected pattern of apparent CSD does obtain in the auditory analysis, and in the articulatory analysis there is evidence of some deletion. Therefore, there is no reason to believe that the use of EMA has precluded the implementation of normal CSD.

3.1.3. Idiosyncratic covert allophony

While only a small portion of inaudible stops exhibited no TT raising whatsoever, there are other relevant facts about the articulatory detail of coronal stop implementation. Specifically, we can conceive of a CSD process whose output is not an articulatory zero, but a discrete category of undershot [t]. Therefore, it is prudent to explore the distributions of individual speakers’ coronal stop TT heights for evidence of multiple categories.

For some speakers, there is a clear unimodal distribution of coronal stop TT heights, with no evidence of discrete categories. Distributions for TT height at the coronal stop target (T) are displayed for these unimodal speakers in figure 3, with separate polygons for audible and inaudible tokens for the reader’s convenience. Here, 0 is the height of the Upper Incisor reference sensor.

While some speakers display unimodal distributions of TT height, with no evidence for discrete categories of target, other speakers show a different pattern. Figure 4 shows distributions of for TT height for the three remaining speakers, whose overall profiles (audible and inaudible combined) are much more bimodal.

In both figure 3 and figure 4, each speaker exhibits a fairly wide range of TT height measurements for audible stops. This might be unexpected given that a coronal stop is canonically made through TT contact at alveolar ridge, which is thought to act as a biomechanical constraint that limits movement.

![Figure 2: TT height trajectory for Speaker 3 producing the sequence ‘striped cat’ in a map task.](image-url)
Figure 3: Raw TT heights at $T$ for unimodal speakers 2 and 5.

with a, quite literal, ceiling effect. Some of the variability in this group of tokens may be due to unavoidable noise in the signal that the Electromagnetic Articulograph records. However, it is also true that the surface on which a coronal closure can potentially be made spans a much larger area than just the alveolar ridge. Indeed, contact with a large portion of the hard palate, the alveolar ridge, or even the upper teeth will produce something that is recognisably a coronal stop. Thus, a speaker could feasibly create a dental closure at the bottom of their incisors that would have a correspondingly low TT height but nonetheless be perfectly audible. Moreover, there is a further possibility that speakers could make laminal closures while producing relatively little TT raising. For the present study, none of these potential issues appear to be fatal. There was a characteristic TT height trajectory in the direction of a closure in almost every case, and no speaker had a majority of audible tokens at the low end of their TT height continuum or a majority of inaudible tokens at the high end.

All of speakers 1, 3 and 5 produce a bimodal distribution of TT heights in their implementation of underlying word-final coronal stops, suggesting that they may have multiple targets. Even if we are reluctant to expand our conception of ‘deletion’ beyond the true articulatory zero, this category of low TT heights must still be explained. The simplest explanation is the idiosyncratic presence of a categorical phonological process, much like the traditional conception of CSD.

There is some variation still within the three bimodal speakers and how audible and inaudible tokens are distributed between each of their two peaks. Speaker 3 exhibits the cleanest divide such that almost all of their lower peak is comprised of inaudible tokens, and almost all of their higher peak is comprised of audible tokens. This is the basic pattern that we might expect under the assumption that the required movement for a coronal stop is raising along an inferior-superior axis and that insufficient raising will not result in a closure (and therefore be inaudible). The picture for speakers 1 and 3 is a little more complicated. Speaker 1 also has two clear peaks but the lower category is populated with several audible tokens as well as most of the inaudible ones. This suggests that a CSD-like process could have an articulatory output – a low TT height – that is non-deterministically correlated with
acoustic and auditory categories. As such, perhaps traditional CSD analyses have been indirectly approximating these categories. Similarly, all three speakers (but especially speaker 4) produce several inaudible tokens with very high TT heights. This is consistent with Browman and Goldstein’s (1990) analysis that stops can be rendered inaudible by the temporal overlap of closures for adjacent segments and nonetheless be fully articulated.

3.2. Is there systematicity in the magnitude of coronal stop lenition?

The previous section served to evaluate the potential for a CSD process as traditionally conceived in light of potential articulatory evidence for and against a widespread process of that kind. We can also evaluate the data in terms of a gradient measure of proportional tongue tip raising and examine systematicity within this dimension. Table 3 shows the fixed effects for a mixed effects linear regression model predicting magnitude of TT raising, with a random for log-interval duration by speaker and a random intercept for word.

Table 3 shows a number of significant effects on magnitude of TT raising. A particularly large effect shows that the log-transformed duration of the interval between points A and B (log.int) affects TT raising such that a longer interval results in higher raising towards a speaker-specific maximum height. The significance of this effect is greatly diminished by the inclusion of a random slope of interval duration by Speaker, allowing for variation in terms of speakers’ physiological capacity to produce articulatory movements in different amounts of time.

There are also effects of task, phonetic environment, and morphological class. Speakers produced significantly less TT raising in the Wordlist reading compared to the Map Task, more TT raising following a dorsal segment compared to a labial segment, and less TT raising in monomorphemes
Estimate Std. Error DF t value Pr(>|t|)
(Intercept) 0.8653 0.1087 6.01 7.958 2.31e-04 ***
zipf 0.0040 0.0052 12.92 0.768 0.456
\textit{t.d} - [t] 0.0178 0.0405 43.53 0.291 0.773
log.int 0.3091 0.0766 4.45 4.035 0.013 *

\begin{tabular}{llllll}
\hline
 & Estimate & Std. Error & DF & t value & Pr(>|t|) \\
\hline
\textit{task} - Script & -0.0669 & 0.0366 & 75.37 & -1.829 & 0.071 \\
\textit{task} - SemDiff & -0.0857 & 0.0450 & 49.59 & -1.903 & 0.337 \\
\textit{task} - Wordlist & -0.3021 & 0.0497 & 44.85 & -6.078 & 2.42e-07 *** \\
\textit{preplace} - Dors & 0.0954 & 0.0291 & 27.49 & 3.277 & 0.003 ** \\
\textit{postplace} - Dors & -0.0148 & 0.0531 & 28.06 & -0.278 & 0.783 \\
\textit{postplace} - Open & 0.0783 & 0.0499 & 151.7 & 1.569 & 0.119 \\
\textit{preman} - Son & 0.0008 & 0.0476 & 27.35 & 0.017 & 0.987 \\
\textit{postman} - Vwl & -0.0054 & 0.0450 & 242.6 & -0.120 & 0.905 \\
\textit{postman} - Paus & 0.0586 & 0.0473 & 271.5 & -1.237 & 0.217 \\
\hline
\textit{gram} - Mono & -0.0680 & 0.0325 & 23.78 & -2.091 & 0.047 * \\
\textit{gram} - Semi & 0.0181 & 0.0648 & 91.77 & 0.273 & 0.786 \\
\hline
\end{tabular}

Table 3: Fixed effects of mixed effects linear regression predicting magnitude of TT raising.

compared to complex words. Some factors that, unexpectedly, did not yield significant results are lexical frequency (\textit{zipf}) and whether the token in question is canonically realised as [t] or [d] (\textit{t.d}).

3.2.1. Articulatory interval duration

The largest effect observed on the magnitude of TT raising is that of articulatory interval duration. This is the time in seconds that elapses between point \textit{A} (immediately preceding \textit{T}) and point \textit{B} (immediately following \textit{T}). The roles of points \textit{A} and \textit{B} are illustrated in figure 1. The simple correlation between interval duration, log-transformed, and TT raising in demonstrated in figure 5, with hollow points for inaudible tokens. The dotted line that intercepts the y-axis at 0 indicates the threshold below which tokens were considered articulatory zeroes.

When not controlling for random slopes by speaker, this effect of interval duration is extremely strong, and it remains whether or not the measure is log-transformed. The data are also nicely, partitioned such that there is a threshold below which no token is audible and above which inaudible tokens are in the minority. This effect is not unexpected. Indeed, it is very reminiscent of classic undershoot effects observed by Lindblom (1963) for variation in vowel quality, whereby shorter vowels were more centralised and less peripheral. It should be noticed that there is a less clear correlation for the absolute values of TT raising. That is, tokens in which the TT was substantially lowered from the baseline at line \textit{AB} have some of the shortest articulatory intervals. We could interpret the effect of interval duration as a phenomenon in the domain of phonetics whereby speech rate conditions TT raising and when speakers allot less time to articulate a coronal stop, they achieve less raising relative to their maximum TT height. However, this does not necessarily constitute evidence against a phonological deletion process. There is no reason that speakers planning to execute a different articulatory target that requires less TT raising should not allot this target a shorter interval duration. When we observe the patterns of TT raising across interval durations for each speaker separately, as in figure 6, there are some interesting findings in terms of this issue of the phonetics-phonology interface and how categoricity can be a diagnostic tool.

Figure 6 displays magnitudes of TT raising, and is not a recreation of the distributions in figures 3 and 4. For this reason, it is extremely noteworthy that we still observe two clusters of tokens in the TT height raising profiles of speakers 1 and 3. Again, the speaker 3 exhibits an extremely clean division of clusters where only one audible token is a member of their cluster with lower TT raising. Speaker 1’s clusters are both populated with several audible tokens, but the bottom end of this speaker’s cluster of tokens with lower TT raising values is comprised of only inaudible tokens. Also interesting to note
Figure 5: Magnitude of TT raising by log-transformed interval duration.

Figure 6: Magnitude of TT raising by log-transformed interval duration for each speaker.
is that speaker 2, whose distribution of raw TT heights was remarkably unimodal, exhibits a small cluster of inaudible tokens in which the TT has lowered a substantial distance. Even speaker 5, the other unimodal participant in terms of raw TT height, produces a couple of inaudible tokens in which the TT raising from baseline $AB$ is negligible.

The results from speakers 1 and 3 reinforce the idea that some speakers may have something resembling a categorical CSD process, with outputs that lie primarily in the articulatory domain, but that are indirectly and non-deterministically perceived in acoustic analyses. Speaker 4, who exhibited the third bimodal raw TT height distribution, presents no discernible pattern in terms of TT raising. However, this speaker also produced the least data. The clusters of inaudible tokens with low TT raising (or even lowering) in speaker 5 and especially speaker 2 suggest that there may also be idiosyncracies in representation at play. In other words, it could be that some speakers store articulatory targets that correspond to segmental information and that they attempt to achieve in real-time, while others may store vectors along which articulators are to be moved. Under this analysis, several of speaker 2’s tokens could be considered to have undergone some process of categorical CSD in that they feature TT lowering rather than raising, even though they are still quite high in terms of raw TT height.

3.2.2. Task

One of the more unexpected results from the regression model summarised in table 3 is that of Task. When considering how speakers may have behaved differently in terms of magnitude of TT raising across the various tasks that were designed to elicit speech, we find that speakers produced the least TT raising in the Wordlist task, and that this was significantly different from the Map task, which contributes to the intercept of the regression model. The distributions of TT raising across each task, collapsed across speakers, is presented in a boxplot in figure 7.

This effect is somewhat surprising. Under the classic view in which different tasks should prompt different degrees of self-monitoring Labov (1972), the word-list is expected to inspire the greatest amount of metalinguistic awareness. As such, we might expect speakers to most noticeably eschew features of casual speech and favour careful and precise speech in this context. CSD, and especially the gradient measure of TT raising used in this study, are excellent examples of a lenition-type phenomenon where lenition can be very narrowly construed in terms of the magnitude of articulatory movement towards a canonical target. Further, Eckert (2003) and Podesva et al. (2006) attribute some prestige and formality to fully articulated and audible coronal stops such that they index social meanings of a high level of competence and education. Therefore, TT raising is a prime candidate as a variable where we would expect style shifting and the highest degree of TT raising in the Wordlist context.

One potential explanation for this effect is somewhat ‘functional’ in nature (Kiparsky, 1972). That is that out of all the tasks there is the least pressure to communicate the stimuli in the Wordlist context clearly to an interlocutor. In all of the other tasks, the participant read or spontaneously produced words in sentences with a researcher listening in. The Wordlist task is the only one in which participants produced words in isolation. However, if we are to appeal to such an ‘information theoretic’ analysis, it remains to be explained why the Map task does not also significantly outperform the Semantic Differential and Script reading tasks in terms of TT raising. Conversely to the situation for the Wordlist, the Map task is the only context in which the speaker is explicitly giving instructions to the researcher to follow. Therefore, we might expect the greatest amount of pressure to communicate clearly in this task, which is not observed. Figure 8 shows magnitude of TT raising by interval duration, with points categorised by task.

One interesting observation from figure 8 tokens in the Wordlist task were produced with the longest interval and the lowest TT raising. The overall trend, as demonstrated figure 5, is for tokens with longer intervals to have more TT raising. This direction is maintained between the ellipses for the Map task, the Script Reading task and the Semantic Differential task, but not the Wordlist. Some potential explanations that take this shape into account are a prosodic explanation and a fatigue explanation. The prosodic explanation is that in the Wordlist participants were required to produce the same word in isolation three times. This means that these words could be considered to all have a strong following phrasal boundary. In addition, each group of three words tended to form an intonational
Figure 7: Distributions for magnitude of TT raising in tokens from each task.

Figure 8: TT raising by interval duration and task, ellipses for 40% confidence interval.
contour across which we might expect a gradual weakening effect. This accounts for the fact that the wordlist is longer and lower because speakers will tend to slow down across an utterance. An even more basic explanation is that participants may experience increasing fatigue across the duration of the experiment. Since the Wordlist task was always conducted last, participants are likely to be at their most fatigued at this point. Therefore, it could be that speakers simply expend less effort to produce coronal stops as they become tired. Teasing apart these competing explanations is beyond the scope of the present study.

3.2.3. Morphological class

One of the most interesting results from the linear regression on magnitude of TT raising concerns the morphological class of tokens. We observe that speakers produce significantly less TT raising for coronal stops at the end of monomorphemic words as compared to coronal stops that constitute an -ed suffix at the end of passive or preterite forms. The distributions for TT raising magnitudes in each morphological class are displayed in a boxplot in figure 9, with significance levels from a basic two sample t-test.

![Figure 9: Distributions for magnitude of TT raising in tokens from each grammatical class.](image)

The results in figure 9 are especially interesting because they are in the same direction of the robustly attested effect of morphological class on CSD. That is, we expect a higher rate of CSD in monomorphemes than complex words, and less TT raising corresponds to less articulatory movement towards a canonical coronal stop closure. This corroborates a similar finding from Purse and Turk (2016), who observe less TT raising for coronal stops in monomorphemes than complex words for a subset of their data: specifically, the speakers of a Southern Standard British English dialect. However, these results pose a problem for the idea that speech production should be strictly modular and that morphology and phonetics should not share an interface. TT raising is a gradient phonetic measure that we do not expect to be conditioned by morphological variables without the mediation of categorical phonology.
4. Pilot Discussion and Proposed Research

The pilot study in this proposal revealed a number of interesting findings. The most convincing is that it is quite common for apparent cases of CSD to involve covert tongue tip raising, as Browman and Goldstein (1990) observed for a handful of cases. However, not all such tokens have high tongue tip raising and some even look categorically deleted (no raising at all). Moreover, we cannot come to firm conclusions on a number of questions about the robustness of various effects and their implications for the representation of CSD, in large part due to the paucity of this type of data:

1. Does categorical CSD exist?
   (a) True zeroes
   (b) Covertly articulated, but discrete, allophones
2. Is categorical CSD (probabilistically) systematic
   (a) Phonological context
   (b) Morphological class
3. Are there gradient effects in the articulation?
   (a) Duration
   (b) Phonological context
   (c) Morphological class
   (d) Task
4. Can we represent CSD in a way that is consistent with previous work and new findings?
5. How sensitive are listeners to the articulatory facts?
   (a) Covert articulation versus true zero
   (b) Do perceptual effects mirror attested production effects?
6. How do individuals differ across all these parameters?

4.1. EMA Corpus

Owing to the potential richness of articulatory data, which is apparent from the pilot study presented here, a large portion of this dissertation will consist of the scaling up of the EMA methodology already implemented for the pilot speakers. My goal for this part of the dissertation is to reach a corpus 40 speakers, including those that have been collected so far. Since each oral EMA sensor must be sterilised, coated in latex, and left to dry overnight, it is only possible to use each of these for one participant per day. For this reason, I think it prudent to devote a whole semester, Spring 2020, to the gathering of this articulatory data.

4.1.1. Categorical CSD

A key advantage of the articulatory data is that it allows us to directly evaluate claims about the representation of CSD in production. It is already evident in the pilot data that CSD with no tongue tip raising is quite rare, contrary to what is suggested in the sociolinguistics canon. However, it is important to verify this finding with a sample of speakers that better compares to the size of traditional variationist studies. This will allow us to better estimate what proportion of inaudible tokens actually constitute ‘true zeroes’, and what factors make these ‘true zeroes’ more likely to occur.

We can also make predictions about the distribution of categorical CSD, that we do not yet have sufficient data to evaluate. One such prediction follows from Guy’s (1991) account of the effect of morphological structure on CSD, which involves a variable CSD process at each of the many levels of a stratified morphophonology. Monomorphemes, which are fully formed from the beginning of this stage of derivation, are eligible to undergo each of these processes, significantly increasing the likelihood that one such process will result in the deletion of a relevant coronal stop. Complex forms, on the other hand, only receive the Level II -ed suffix at the final level of a stratified morphophonology. Therefore, coronal stops in complex words are only the potential target of one variable CSD process and are far less likely to be deleted.
Bermúdez-Otero (2010) and Myers (1995) draw from Guy’s (1991) account, explaining that apparent deletion in monomorphemes should be comprised of far more instances of categorical deletion than in complex forms, since the former type of word should be eligible for a greater number of potential applications of a phonological CSD process compared to the latter. Myers (1995) in particular invokes Zsiga’s (1993) ‘Lexical-Categorical Hypothesis’, that the output of lexical phonology must be categorical and cannot be gradient. This implies a relaxed assumption of categoricity at the postlexical level, where a final phonological CSD process may target monomorphemes and complex forms alike. Such a position is informed by an assumption that the phonology must operate on strings of discrete symbols, and so levels of phonology whose output becomes the input to a subsequent level of phonology cannot have a gradient output. The perspective that postlexical phonology may not have a categorical output makes an even stronger prediction with regard to the distribution of categorical CSD across morphological classes, because if the postlexical CSD process is gradient complex forms may not be subject to any categorical phonological CSD process at all. Thus, both Bermúdez-Otero (2010) and Myers (1995) expect more categorical CSD in monomorphemes than complex forms.

On the other hand, Tamminga’s (2016) work makes a different prediction based on evidence from persistence, an effect of naturalistic priming whereby a speaker is more likely to apply a variable process that they have just applied immediately beforehand. She finds that apparent CSD in monomorphemes (e.g. pact) primes CSD in the exact same word (pact), but not in a different monomorpheme (e.g. soft) nor in a complex word (e.g. packed). However, apparent CSD in a complex word (e.g. packed) primes CSD in the same word (packed) and in other complex words (e.g. cracked). From this, Tamminga (2016) surmises that at least some apparent CSD in complex words should be attributed to zero-allomorphy. The selection of this zero-allomorph in the place of the canonical -ed suffix then makes its subsequent selection for another complex form more likely. Since a zero-allomorph will have entered the phonological derivation without phonological content, it should necessarily present itself as an instance of categorical deletion. Therefore, zero-allomorphy provides an avenue not available to monomorphemes through which complex forms could look at though they have undergone categorical CSD.

Interestingly, the predictions from Myers (1995) and Bermúdez-Otero (2010) on one hand, and Tamminga (2016) on the other, are not necessarily mutually exclusive. Both variable allomorphy and lexical phonology could be sources of what looks like categorical CSD. However, the output of these processes may not be identical. Under an assumption that phonology is not constrained to be ‘structure-preserving’, a coronal stop token with residual TT raising is compatible with having undergone a categorical process in the phonology. However, a zero-allomorph must receive its phonological form from the underlying inventory, and cannot then be specially targeted for a phonological process. Therefore, zero-allomorphy must be ‘structure preserving’ and should not involve any TT raising towards a canonical coronal stop target. Saliently, there is evidence in the pilot data presented here for both true zeroes (≤0 TT raising) and tokens with TT raising of a small enough magnitude to perhaps be categorically distinct from a fully realised stop. This latter token type could be consistent with a structure non-preserving CSD analysis.

4.1.2. Gradient articulatory effects

In addition to shedding light on the question of categoricity in the production of CSD, the pilot data shows some interesting results in the gradient measure of TT raising magnitude. Namely, that this value varies systematically according to a number of factors. These factors include the interval duration, phonetic environment, morphological class, and task. This suggests that, although CSD has traditionally always been described in categorical terms, there may instead be an equally important gradient phonetic component that contributes to patterns of apparent CSD in the acoustic signal. The scaling-up of the articulatory data is will enable us to test the robustness of these gradient effects. Moreover, a larger dataset will allow for the exclusion of more datapoints to properly determine whether the observed effects are indeed gradient and not, for example, a function of a token’s likelihood to be a member of one or another discrete clusters in TT raising magnitude. A truly gradient effect should
be observed within one or more contiguous clusters on tokens, e.g. among the tokens with pronounced TT raising.

The observed effect of morphological class on TT raising is particularly interesting. We observe that a larger magnitude of TT raising occurs for the stops at the end of regular past forms than for stops at the end of monomorphemes. This resembles the well-attested effect of morphological class on CSD—higher rates of CSD in monomorphemes than complex forms—in that less TT raising can be interpreted as a less complete articulation of a coronal stop, i.e. closer to zero on a continuum from zero to full [t]. However, a strictly modular view of speech production stipulates that an effect of morphology should have its reflex in the categorical phonology, not the gradient phonetics. This effect in particular, and the relative paucity of evidence for categorical CSD, point to the need to explore alternatives for the representation of the phenomenon.

4.2. Prosody study

One way to approach the representation of CSD that allows for both morphological conditioning and residual TT raising is to explore ways in which differences in morphological structure might be preserved in the phonology. Evidence of this might be found in something like Incomplete Neutralisation, in which underlying constraints are reduced, but not completely lost as has traditionally been assumed. Much of the research into Incomplete Neutralisation has focused on ‘devoicing’ phenomena, finding that segments that have undergone ‘devoicing’ still bear phonetic correlates (typically a longer preceding vowel) of greater voicing compared to underlingly voiceless segments (Braver, 2014; Roettger et al., 2014). This is relevant to the morphological effect in CSD, because this effect is confounded with a difference in the underlying coronal stops themselves. The -ed suffix in passive and preterite forms is underlingly voiced and undergoes devoicing when it follows a voiceless segment (e.g. packed [pækt] versus grabbed [gəæbd]). This underlying difference may lead to small phonetic differences that result in different rates of CSD. However, as with all cases of Incomplete Neutralisation, this would be difficult to explain. Either the phonetic implementation has to ‘see’ the underlying form that it came from, or else we must allow structure non-preserving processes to be far more ubiquitous than is the norm in phonological theory.

Along similar lines to the Incomplete Neutralisation story, but less challenging to some fundamental assumptions in phonology, we could turn to Indirect Reference Hypothesis (Inkelas, 1990). This is the idea that morphosyntactic structure can have phonetic correlates without a direct interface between the two because prosodic structure intervenes and preserves important contrasts in the morphosyntax. This is relevant to CSD because some evidence suggests that different prosodic structure could align with morphological structure even at the within-word level. Thus, the purpose of the second component of the dissertation is to explore the viability of the idea that monomorphemic words have a different prosodic structure to morphologically complex words, and this could account for different CSD rates.

4.2.1. The Indirect Reference Hypothesis

Walsh and Parker (1983) find that a Level II suffixal [s] (e.g. wrecks, laps, hearts) is produced about 9 ms longer on average than an [s] at the end of a monomorphemic homophone (Rex, lapse, Hartz). Similarly, Schwarzlose and Bradlow (2001) find that stem-final consonants followed by suffixal [s] (e.g. [k] in tacks, tucks, and macs) were 3 to 5 ms longer than the penultimate segment in monomorphemic homophemes (tax, tux, max), and Sugahara and Turk (2009) report that stem-final rhymes before Level II English suffixes are 4-6% longer than corresponding strings in monomorphemes. A leading explanation for these findings is that any roots and Level I suffixes to which a Level II suffix can be attached form a prosodic constituent preceding said Level II suffix. As such, prosodic lengthening processes target material around the domain-final boundary that accompanies Level II suffixes. If we are to accept this analysis, regular -ed suffixes that indicate passive or preterite forms should also be the site of prosodic lengthening processes. As such, perhaps the effect of morphological class is actually a product of an increased duration, on average, of phonological material around a prosodic boundary that accompanies an -ed suffix.
We already observe a fairly strong correlation between interval duration and TT raising in figure 5, such that a greater interval duration allows for higher TT raising. Placing the locus of this effect in the prosody allows for an effect of morphological structure on phonological representation that has gradient phonetic effects in terms of TT movement.

4.2.2. Prosody study design

To test the hypothesis that there are prosodic lengthening effects around level II suffixes, I will make acoustic recordings of speakers producing synonyms and near-synonyms that differ in terms of morphological structure. These will include but not be limited to words that end in an underlying coronal stop. I will then use Praat to hand-measure durations from the following points that have been identified as important in the literature on this topic:

- The rime of the stem in the complex form and the corresponding string in its homophonous monomorpheme (e.g. [ak] in tacks and tar)
- The suffix in the complex form and the corresponding string in its homophonous monomorpheme (e.g. [t] in pact and packed)

Where previous studies on this kind of lengthening have limited their analysis to just one measure of duration, this will enable me to explore any potential relationship between the two. I will use Sugahara and Turk’s (2009) method of presenting stimuli in a consistent context with a preceding phrase to introduce the target word’s meaning. Some example pairs are found in table 4.

<table>
<thead>
<tr>
<th>Set 1</th>
<th>Set 2</th>
<th>Set 3</th>
</tr>
</thead>
</table>

Table 4: Example words with -s, -ed, and -ing suffixes and their monomorphemic (near) homophones.

The aforementioned methodology will allow for controlled testing of the relevant variables, uninhibited by the EMA sensors used to gather articulatory data. This is valuable. However, as motivated the inclusion of articulatory data in this dissertation in the first place, various details of speech production are not evident from the acoustic signal. Therefore, the data from the wordlist portion of the scaled-up EMA corpus, which will also include several relevant homophone pairs, will also be explored for evidence of prosodic effects. As such, it will be possible to examine the phasing of gestures for these words, not just the timing of the completion of closures that is evident from the acoustic signal.

4.3. Perception study

One large area of inquiry that is complicated by the results of the pilot EMA study concerns how CSD is learned. Previous studies have observed that children exhibit apparent CSD with the same conditioning factors as in adult speech by around age 4, and by age 7 they reliably match the actual rates of CSD produced by their adult caregivers (Labov, 1989). However, based on the pilot study presented in this document, it is likely that not all tokens that researchers have labelled ‘deleted’ are produced the same, and it is not clear to what extent the learner is privy to these articulatory details in the input.

It seems distinctly possible that the input to the learner is comprised of a mixture of tokens that are categorically deleted and tokens that are rendered inaudible through undershoot or articulatory overlap. If learners are unable to distinguish cases like these it is unclear why, for example, the rate of CSD is not rapidly increasing over time. One might predict that a learner who cannot distinguish between cases of apparent CSD would, as researchers have, evaluate all such cases as instances of categorical deletion. This learner would presumably acquire and produce categorical CSD at a rate that matches the rate of apparent CSD in their input, and additionally produce some portion of
undeleted stops that are rendered inaudible by the same phonetic processes of undershoot and overlap, resulting in an increased rate of apparent deletion in the input for the subsequent generation. Instead, it appears that learners’ apparent CSD rates match those of the input. This either means that learners can perceive and directly acquire the different strategies for producing apparent CSD, or else that each learner infers strategies for matching the rates that they observe.

4.3.1. Perception study design

The first step towards understanding how CSD is acquired and represented across time is investigating how it is perceived. To that end, the third component of this dissertation will focus on listeners’ perception of CSD. The stimuli for this experiment will be drawn from the created EMA corpus, so that the articulatory details can be included in the analysis of their perception. Short excerpts of the acoustic recording from the EMA corpus containing critical items in context will be played for online participants. A selection will be taken from each speaker in the EMA corpus. Selections will be balanced, as much as possible, for phonological context, morphological class, and EMA-informed production outcomes (audible, inaudible with TT raising, inaudible with no/very low TT raising). The perception experiment participants will be asked questions of the following form, and asked to rate the excerpt in response:

- ‘How clearly did the speaker produce the [t] at the end of pact’

Ratings will be taken on a scale from 1 to 4 (‘very clear’, ‘somewhat clear’, ‘somewhat unclear’, ‘very unclear’). This question is designed to access speakers’ perception of the surface form in particular, hopefully minimising the effects of inferences or biases based on the underlying form of the target word or its well-formedness with no coronal stop.

This study will aim to accomplish a number of broad goals. First, it will corroborate the impressionistic judgments that I have made for whether or not a [t] is audible. While the vast majority of classic research on CSD makes use of these judgments, it is rare that they are rigorously compared against the perception of multiple listeners. Second, this data will shed light on listeners’ knowledge of /t/ and its potential realisations. In particular, it is important to find out if listeners can distinguish between /t/s that are traditionally labelled as deleted, with and without TT raising.

5. Conclusions

The data in the pilot study presented here do not fully support the view that there is widespread categorical CSD, nor do they reinforce a perspective that all cases of apparent CSD are merely the result of connected speech processes. According to articulatory evidence, categorical CSD as we have defined it is far rarer than previously thought, but still constitutes at least 17% of inaudible coronal stops. Further, the pilot speakers exhibit systematicity in the magnitude of TT raising beyond what is captured by a dichotomy between presence and absence of movement in the direction of an alveolar ridge target.

These results give rise to difficult questions about CSD, and perhaps other variable phenomena whose representation is similarly ambiguous. We still cannot say conclusively whether or not there is some phonological process that resembles the traditional description of CSD; even a rare one. Alternatively, we cannot rule out that cases of missing TT raising may constitute one extreme end on a continuum of lenition. If categorical CSD does exist, we will need to explore the factors that (probabilistically or deterministically) condition it, and to what extent they resemble the factors relevant to apparent CSD. Within any categories that can be identified, it will be important to investigate the possibility of gradient effects on the articulation of coronal stops. Given the results from all of these lines of inquiry, we must reassess how to explain CSD. From the outset, the new representation of the phenomenon will have to account for robust patterns of apparent CSD, a large portion of the instances of which are not categorical. Additionally, it will be necessary to explore listeners’ capacity
to perceive the details of articulation that are observed. Finally, as is already suggested by the pilot data, individual participants may not all exhibit the same patterns in production (or perception). This means that there may not be just one answer for many of the questions I have posed, but we will still be able to consider how behaviour along these dimensions is constrained.

In sum, it remains largely unclear how CSD should be represented, if any one kind of representation can even account for all of the data. However, the results from the pilot study suggest that the more data of this type is likely to shed light on many of the most pressing questions surrounding this issue. Therefore, the core this dissertation will comprise a large-scale articulatory investigation into CSD, supplemented by studies to explore alternatives to how CSD has traditionally been conceived. It will, ultimately, constitute one of the first investigations into the details of the production and perception of CSD, despite a wealth of studies that have assumed the phenomenon to have various properties.

6. Timeline

<table>
<thead>
<tr>
<th>Month/Range</th>
<th>Activities</th>
</tr>
</thead>
<tbody>
<tr>
<td>January 2020</td>
<td>Defend dissertation proposal</td>
</tr>
<tr>
<td></td>
<td>Revise proposal as recommended</td>
</tr>
<tr>
<td>February-May 2020</td>
<td>Recruit and run EMA participants</td>
</tr>
<tr>
<td></td>
<td>Recruit and run prosody study</td>
</tr>
<tr>
<td>June-August 2020</td>
<td>Code EMA data</td>
</tr>
<tr>
<td></td>
<td>Run perception study online</td>
</tr>
<tr>
<td></td>
<td>Write up prosody study</td>
</tr>
<tr>
<td>August-October 2020</td>
<td>Run and code any remaining EMA participants</td>
</tr>
<tr>
<td></td>
<td>Write up perception study</td>
</tr>
<tr>
<td>November-December 2020</td>
<td>Write up EMA study</td>
</tr>
<tr>
<td>January-April 2020</td>
<td>Combine and polish dissertation</td>
</tr>
<tr>
<td>May 2020</td>
<td>Defend dissertation</td>
</tr>
</tbody>
</table>

References


Tamminga, M. (2018), ‘Modulation of the following segment effect on english coronal stop deletion by syntactic boundaries’, Glossa 3(1), 86.


