1. Introduction

In American English, pairs of words such as 'prince' and 'prints' are predicted to be homophonous in surface realization through the epenthesis of a stop in nasal-fricative clusters (in this case, $\emptyset \rightarrow [t] / n \_ s$). In such a case, the difference between underlying and epenthetic [t] is said to be neutralized. However, production studies conflict as to whether the surface realizations of these words are actually identical (Fourakis and Port, 1986; Yoo and Blankenship, 2003; Arvaniti, in sub). The current study, using both closure duration and closure-to-word-length ratio as indicators of the presence of a stop, showed that speakers could reliably produce a difference between underlying and epenthetic stops in nasal-fricative clusters and that this difference was affected by speech style. Considered in the context of previous studies, these results indicate that [t] epenthesis in nasal-fricative clusters is a case of near-merger rather than incomplete neutralization as has been previously proposed (Fourakis and Port, 1986; Arvaniti, in sub). Incomplete neutralization and near-merger are similar in that in both cases there are differences in production between underlying forms that are predicted to be identical in the surface realization but they contrast in the role that those differences, called sub-phonemic cues, play in perception. With incomplete neutralization, the sub-phonemic cues can be used with above chance accuracy by listeners to disambiguate underlyingly contrastive forms (Port and O'Dell, 1985; Warner, 2001) but this is not the case in near-mergers (Diehm and Johonson, 1987; Yu, Janson and Schulman, 1983).

Historically, the key production experiment for stop epenthesis in nasal-fricative clusters
was Fourakis and Port (1986). They showed that stop epenthesis in nasal-fricative clusters occurred in American English but not South African English and that stop epenthesis could not therefore be the result of a universal constraint on articulation as proposed by Ohala (1974). The Fourakis and Port study also showed the inadequacies of a classic phonological analysis to explain the data. A classic phonological neutralization predicts that epenthetic and underlying stops in nasal-fricative clusters would be identical in surface realization. However, Fourakis and Port found that there were in fact measurable differences between underlying and epenthetic [t] in the /ns/ clusters. Fourakis and Port proposed that for American English, [t] epenthesis in /ns/ clusters is a case of incomplete neutralization; a regular variation in the speech signal that cannot be accounted for by phonetics nor by classic phonological models. To account for the data, they proposed phase rules, a set of low-level phonological rules that would govern articulatory transitions between segments.

Almost 20 years later, Yoo and Blankenship (2003) could find no evidence in American English for incomplete neutralization between epenthetic and underlying [t] in /ns/ clusters, but Arvaniti (in sub.) showed that incomplete neutralization occurred but only in unfamiliar words. Based on this evidence and known patterns of sound change over time, Arvaniti suggested that the neutralization had evolved and become more complete since the Fourakis and Port study.

The results of the Arvaniti study indicate that there are at least some contexts in which the neutralization between underlying and epenthetic [t] in /ns/ clusters is still incomplete for American English speakers. One context that has not yet been explored for [t] epenthesis in /ns/ clusters is the effect of speech style. In effect, research on another kind of incomplete neutralization, final devoicing, showed that speakers produced greater differences between underlyingly contrastive forms that were predicted to be surface identical in some speech styles that others (Port and
This study reports on the effects of speech style for the neutralization of [t] in /ns/ clusters in word-final position. The results showed a difference between epenthetic and underlying [t]s in the three speech styles studied but the difference was greater when speakers attempted to disambiguate the words for a researcher who could not see the prompts.

In a broader context, the evidence suggests that although [t] epenthesis in /ns/ clusters is not a case of complete neutralization, it is also different from the incomplete neutralization typified by final devoicing. For final devoicing, the sub-phonemic cues that distinguish segments that are predicted to be surface identical do not seem to be diminishing over time as studies continue to show robust evidence of incomplete neutralization (Port and O'Dell, 1985; Port and Crawford, 1989, Warner, 2004). For final devoicing, coarticulatory cues of underlying voicing persist to some degree in the surface realization (Port and O'Dell, 1985; Port and Crawford, 1989) but such coarticulatory cues have not been found for [t] epenthesis in /ns/ clusters since the original Fourakis and Port study (1986). Finally, perception studies of final devoicing have shown that listeners can use sub-phonemic cues to disambiguate underlingly contrastive forms (Port and O'Dell, 1985, Port and Crawford, 1989, Warner et al. 2004) but such evidence has proven elusive for [t] epenthesis in /ns/ clusters (Shinya, 2005, Arvaniti, in sub.).

This disjunction between production and perception of epenthetic [t] in /ns/ clusters where acoustic cues are reliably produced but do not seem to encode linguistic information is not unique. Such a disjunction has been described for Russian palatalized consonants (Diehm and Johnson, 1997), Swedish vowels (Janson and Schulman, 1983), 'Near' and 'Square' diphthongs in New Zealand English (Hay et al. 2006) and Cantonese lexical and morphologically derived mid-rising
tone (Yu, 2007) and is called a near-merger.

2. Background

Phonology and phonetics are two related areas of language that are often described as separate from each other. According to phonological models, phonemes are abstract and categorical and interactions between segments can be described in terms of phonological rules or constraints. Phonetics describes the actual acoustic signal or articulation in the context of speech production and perception.

In some circumstances, there may be a regular difference between the underlying abstract phonological form and the surface realization. In classic models, there are two methods arriving at an underlying-surface difference. The first is the implementation of a phonological rule. In this case, the underlying form cannot be recovered from the surface form. For example, in a case of nasal assimilation where the nasal becomes homorganic with the following segment as shown in (1), we would not be able to tell the underlying form from the surface realization. The /n/ produced as [ŋ] in the surface realization would be indistinguishable from [ŋ] that was underlyingly /ŋ/.

(1) "incredible"  
/in + kredibɔl/  
[ŋkredibɔl]  

<p>| | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>&quot;incredible&quot;</td>
<td>gloss</td>
</tr>
<tr>
<td></td>
<td>/in + kredibɔl/</td>
<td>UR</td>
</tr>
<tr>
<td></td>
<td>[ŋkredibɔl]</td>
<td>SR</td>
</tr>
</tbody>
</table>

The second type of difference is the consequence of coarticulatory constraints operating in

---

1. Assuming that phonetic implementation only has access to the surface form of the phonology.
the phonetics. In these cases, variation is caused by either constraints linked to the actual physiology and aerodynamics of the vocal tract, or interactions between segments produced in sequence. In these cases, we would expect the variation to hold across and within languages. For example, a non-nasal vowel produced before a nasal consonant will become at least partially nasalized. The velum is raised in anticipation of the nasal during the vowel and this creates a noticeable nasalization of the vowel. This is a mechanical constraint and therefore holds across languages (Ohala, 1975).

There is, however, evidence for a third type of underlying-surface difference that combines factors distinguishing it from the two preceding forms. This third type of variation is not constant across and between languages and the underlying form is recoverable from the surface realization. In short, although the phonology of a language predicts that two sounds will be neutralized in surface realization as for the example in (1), that change is acoustically incomplete; the resulting segment is different both from its underlying form and from the category that it was predicted to become. The existence of this type of variation, called incomplete neutralization, and the reasons for which it might exist, are important subjects of research that inform us about the nature of the phonology-phonetics interface and the role of sub-phonemic cues in speech perception.

The most well-attested and studied case of incomplete neutralization is that of final devoicing. Reviewing the research on final devoicing is helpful in showing the range of variables that need to be evaluated in the study of incomplete neutralization.

Final devoicing occurs in a variety of languages such as Russian, Dutch, German and Polish and is classically described as the result of a phonological rule that neutralizes the contrast in voicing by devoicing word-final voiced stops as shown in an example for Dutch in (2) (example
from Warner et al. 2004).

(2)  

<table>
<thead>
<tr>
<th>underlying</th>
<th>surface</th>
<th>orthography</th>
<th>gloss</th>
</tr>
</thead>
<tbody>
<tr>
<td>/rat/</td>
<td>[rat]</td>
<td>rat</td>
<td>'rat'</td>
</tr>
<tr>
<td>/rad/</td>
<td>[rat]</td>
<td>rad</td>
<td>'wheel'</td>
</tr>
</tbody>
</table>

However, a number of production studies (Port and O'Dell, 1985, Port and Crawford, 1989, Warner et al. 2004 and others) have shown that there is a difference in the realization of these forms at the phonetic level. The duration of the voiceless closure may be shorter for underlyingly voiced stops than for underlying voiceless stops. Other coarticulatory cues to the underlying voicing of the final consonant, such as vowel length and burst duration, may subsist. For example, in general the vowel that precedes a voiced consonant is longer than one which precedes a voiceless consonant. After final devoicing has operated, the vowel preceding the underlyingly voiced consonant is longer than that of an underlyingly voiceless consonant. Perception studies have shown that these sub-phonemic cues can be used by listeners at above chance rates for disambiguation of forms that particular phonological analyses have predicted to be homophous (Port and O'Dell, 1985; Port and Crawford, 1989, Warner et al. 2004).

Environment and type of segment may also affect the level of neutralization. Charles-Luce (1985) showed that the degree of neutralization was affected by the position of the word in the phrase and the manner of articulation of the final consonant. The difference in voiceless closure as a function of underlying voicing was only significant in clause-final stops and the difference in duration of the preceding vowel was only significant in clause-final fricatives.

Orthography also influences incomplete neutralization as the differences between underlyingly contrastive consonants are smaller when the experimental methodology does not
emphasize spelling. In an experiment on German, Fourakis and Iverson (1984), elicited conjugated verb forms based on the base verb form (as an English speaker might answer “wake, woke, woken” in response to the prompt “to wake”). In this condition, Fourakis and Iverson found no differences between underlyingly voiced and voiceless consonants. Port and Crawford (1989), again for German, asked listeners to repeat a sentence recited by a researcher. Although speakers did produce a difference between underlyingly contrastive forms, it was much smaller than in conditions where they actually saw the words written down.

There is also evidence to suggest that context and speech rate play an important role in incomplete neutralization. Port and Crawford (1989) used a methodology that incorporated different speech styles and levels of ambiguity and concluded that speakers could modulate the degree of neutralization. They found that speakers produced greater levels of contrast between minimal pairs contrasting the underlying voicing of the final consonant when they were asked to dictate test words to a researcher than when they recited a word list, or repeated phrases produced by a researcher.

The case of incomplete neutralization under consideration here is that of the epenthetic stops in nasal-fricative clusters that occur in some varieties of English including Southern California American English, the variety investigated here. This process is responsible for words such as 'prince' being pronounced as [prınts]. The epenthetic stop that appears in nasal-fricative clusters is described by Anderson (1975) as sharing place of articulation with the preceding nasal and voicing with the following fricative.
Ohala (1974) proposed that rather than being the realization of a phonological rule, stop epentheses are consequences of universal articulatory constraints. Concerning the epentheses of stops in nasal-fricative clusters (for minimal pairs such as dense-dents), Ohala suggested that if the velum is raised before the release of the nasal, intra-oral pressure would increase. It would be the release of this pressure as a burst that listeners mistakenly perceive as a [t].

In an aerodynamic study of stops in nasal-fricative clusters, Ali, Daniloff and Hammarberg (1979) proposed a slightly different mechanism for production of epenthetic stops. Their airflow experiment showed a cessation of voicing before the end of the nasal corresponding to a period of
silence in the wave form. Even though air was still being released nasally, the intra-oral pressure rose as the airflow in the supra-glottal tract increased due to the cessation of voicing. The intra-oral pressure was released as a burst at the beginning of the fricative. Ali et al. suggest that this burst, combined with the silence of the voiceless nasal, could be perceived as a stop.

Historically, the key production experiment in the area of stop epenthesis is Fourakis and Port (1986) because it showed that [t] epenthesis in /ns/ clusters could not be the result of universal articulatory constraints as proposed by Ohala. Fourakis and Port used both real and nonsense monosyllabic words that contrasted the elements of the word-final sonorant-fricative cluster: the sonorant was [n] or [l], the underlying stop was present or absent and the final fricative was voiced or unvoiced. They defined a stop as an interval of silence greater than 10ms with no visible acoustic noise in the spectrogram between the sonorant and the fricative. The results showed a difference between speakers of South African English and American English. The 4 speakers of South African English did not produce stops in nasal-fricative clusters where there was no underlying stop, and the five American speakers did. This variation is not compatible with a purely articulatory explanation predicting that the rate of occurrence of epenthetic stops, given the right context, would be constant across languages and dialects. Fourakis and Port also found that epenthetic stops, when they occurred, were more frequent following a nasal and preceding a voiceless fricative than in other environments.

Fourakis and Port (1986) also gave the first empirical evidence that stop epenthesis is a case of incomplete neutralization by showing that the closures for epenthetic [t]s produced by American English speakers were actually shorter than those for underlying [t]s. They also found some coarticulatory differences. For example, in the environment of an underlying [t], the preceding
vowel was shorter than for epenthetic [t]. They suggested that underlying [t] was planned and therefore the length of surrounding segments was adjusted to its presence. For epenthetic [t], which assumed to be unplanned, there is no such modification in the environment; the length of other segments is conditioned not on the presence of epenthetic [t], but on its absence.

The goal of subsequent production studies of stop epenthesis in /ns/ clusters has been to reproduce the clear distinction between underlying and epenthetic stops observed by Fourakis and Port and to more narrowly define the environments in which an epenthetic [t] is likely to occur.

Yoo and Blankenship (2003) evaluated the role of stress and word position on the occurrence and duration of epenthetic and underlying [t] in /ns/ sequences in a production study. They compared clusters in word-medial and word-final position in two stress conditions, after a stressed vowel and after an unstressed vowel, in the productions of seven speakers. Measurements were performed in a manner analogous to the Fourakis and Port study.

(5) Examples of conditions and words used by Yoo and Blankenship.

a. word-medial after a stressed vowel  
   'census' – 'dents us'

b. word-final after a stressed vowel  
   'intense' – 'in tents'

c. word-medial after an unstressed vowel  
   'consent' - 'blunt-saying'

d. word-final after an unstressed vowel  
   'science' – 'consents'

The results showed that epenthesis occurred more frequently in word-final clusters than in word-medial clusters. However, there was no evidence for incomplete neutralization as when they occurred, epenthetic [t]s were no shorter than underlying [t]s.

Yoo and Blankenship (2003) also evaluated the TIMIT corpus using the same criteria of measurement and word types as for their production experiment. The TIMIT corpus is a set of 6300
utterances in American English produced in laboratory conditions, segmented automatically and then manually corrected. The results of this corpus study confirmed the findings of the production experiment concerning [t] occurrence as a function of word position but the corpus data did show a significant difference in duration between epenthetic and underlying [t]. Yoo and Blankenship did not have an explanation for the divergence of results between the production study and the corpus study.

One issue with the both the Fourakis and Port study (1986) and the Yoo and Blankenship study (2003) study is that they evaluate the presence of a [t] based solely on the presence of an interval of silence in the spectrogram. However, perception and airflow studies show that bursts may also play an important role.

In the articulatory account of epenthetic stops proposed by Ali et al. (1979), an epenthetic stop is characterized by a build-up of intra-oral pressure after cessation of voicing for the nasal that produces a burst when released. Perception studies (Warner et al., 2001, Shinya, 2005) have shown that both burst and silent interval can trigger the perception of a stop in Dutch and American English respectively. In the most recent and comprehensive production study to date, Arvaniti (in sub.) addressed this concern by evaluating duration and occurrence for both closure (silent interval) and burst.

Arvaniti hypothesized that the incomplete neutralization of epenthetic stops in American English is tending towards completeness and that this would explain the disparity of results between the Fourakis and Port (1986) study and the Yoo and Blankenship (2003) study almost 20 years later. Based on this, Arvaniti predicted that the neutralization would be less complete in infrequent and unfamiliar words and in the productions of older speakers. Arvaniti evaluated the
production of epenthetic [t] in /ns/ clusters in word-medial and word-final position using test words that varied in frequency and familiarity by Southern California English speakers from two age groups (18-23 year-olds and 40-55 year-olds).

She found differences between epenthetic and underlying stops but only in unfamiliar words and in word-medial position. Arvaniti suggested that this was compatible with patterns of sound change over time; the contrast between underlying and epenthetic [t] was maintained in environments where speakers encounter epenthetic [t] more rarely.

It seems likely that Arvaniti was correct in assuming that the neutralization is evolving towards completeness, at least in the context of word final /ns/ clusters for Southern California English speakers, as no study since Fourakis and Port (1986) has shown evidence for a difference between epenthetic and underlying stops in the word-final /ns/ clusters where they most frequently occur.

However, one dimension of [t] epenthesis production has not yet been explored. In previous production studies, speakers produced test words in either a naturalistic phrase (Arvaniti, in sub.) or a laboratory speech style carrier phrase (Fourakis and Port, 1986; Yoo and Blankenship, 2003). It is possible that speakers only produce a contrast when the context requires them to do so. Port and Crawford (1989) showed for final devoicing that when speakers were in a situation where they were forced to produce a contrast between underlyingly voiced and voiceless consonants, there was a greater difference in the surface realization than in other contexts. The speakers in the Port and Crawford study were asked to dictate a phrase containing a minimal pair contrasting underlyingly voiced and voiceless segments in such a way that a person who could not see the prompt would know which of the two words was first.
Based on the results of the Port and Crawford study, the purpose of the current experiment was to discover whether speakers could produce a contrast between underlying and epenthetic stops as a function of speech style. Productions of words contrasting in word-final /ns/-/nts/ clusters were elicited in three different speech styles, a naturalistic phrase, a lab style list and a forced contrast phrase similar to that of the Port and Crawford study. Based on the hypothesis that the neutralization between underlying and epenthetic stops in nasal-fricative clusters is nearly complete and that speakers would produce greater differences between underlying and epenthetic stops when the context required it, the following hypotheses were formulated:

- In the Natural and List styles, there would be durational differences between underlying and epenthetic [t] only in infrequent and unfamiliar words.
- In the Contrast style, there would be a clear difference in duration between epenthetic and underlying [t].
- Epenthetic [t] would occur more often in frequent and familiar words in all conditions.

3. Methods

3.1 Materials

Three minimal pairs of words, *mince-mints, prince-prints and quince-quints*, contrasting underlying and epenthetic stops in word final /ns/ clusters were selected. These three pairs represented three levels of frequency and familiarity. Frequency and familiarity levels were based on the WU Speech and Hearing Lab Neighborhood Database and confirmed by ratings on a 7 point scale (with 7 =
most familiar) by 16 undergraduate students at UCSD not screened for language background. *Prince* and *prints* were chosen because they are both familiar and frequent. *Quince* and *quints* are both unfamiliar and infrequent. However, *mince* is familiar and infrequent while *mints* is familiar and frequent so the use of this pair could help separate familiarity from frequency effects.

The words were presented in three speech style conditions: Natural style, List style and Contrast style. In the Natural style, the word was contained in a naturalistic phrase such as "Jane saw the handsome prince, but she didn't like him" that did not draw attention to the test word. In this condition, the test words always appeared before a phrase break to control for sentence level prosody. The second speech style was the List condition. In this case, the words were presented in a classic lab-speech style carrier phrase: "I will say *prince* a few more times". The final condition was Contrast. It was designed to force the speaker to produce a contrast between the words in a minimal pair: "I said *prince* not *prints*" and "I said *prints* not *prince*". Fillers were created for both the Natural and List speech styles. The purpose of the filler sentences was to obscure which words were being investigated. The fillers were sentences of the same format as those containing the test words but with the test words replaced by words unconnected to the experiment. In the Contrast style, as the speakers were aware of the identity of the test words, no fillers were used.

5. Carrier phrases used for the Natural Condition

- Jane met the handsome prince, but she didn't like him.
- I used to buy a lot of prints, but they tend to fade.
- I really like mints, but I almost never eat them.
- The British call it mince, but we call it ground beef.
- There used to be no quints, but fertility drugs have changed that.
- There is a fruit called quince, but almost nobody knows what it is.
3.2 Speakers

The participants were 12 native speakers of Southern California English aged between 18 and 25 years old. They were recruited from UCSD undergraduate classes in Linguistics and received extra credit in those classes in exchange for their participation. None of them were fluent in another language or had parents that were not native speakers of English. No hearing or speaking disorders were reported. None of the participants were aware of the purpose of the experiment.

3.3 Procedures

In the Natural and List speech styles, the carrier phrases and fillers were produced seven times in pseudo-random order. This randomization had two functions. It ensured an even distribution between the fillers and test words and varied the context of surrounding sentences in which each sentence was produced. The stimuli were presented one sentence at a time in a PowerPoint presentation. The pace of the presentation was controlled by the speaker. The three conditions were presented in the following order: Natural, List, Contrast. The Natural condition was presented first so that speakers would be as unaware as possible of which words in the sentence were being tested. For the Natural condition, the speakers were asked to read the sentences as naturally as possible. In the List condition, the speakers were asked to produce the sentence more carefully. The Contrast condition ('I said prince not prints' and 'I said prints not prince') was presented last because it required that the speakers be made specifically aware of the minimal pairs. For this condition the researcher sat with her back to the speaker who was told that the researcher would be writing down the underlined word in each sentence and that it was the speaker's task to produce the sentence in
such a way that the researcher would be able to guess what the underlined word was (only the underlined word was used in measurements). Speakers were asked to repeat any sentences that were produced with any dysfluency. They could take a break of up to 2 minutes at the end of each condition.

Speakers were recorded in a Model 40A5-T1 audiometric booth with an Aiwa DM-H200 cardioid dynamic microphone using an AD converter sampling at 44.1KHz. and the Wavepad software on a PC. The researcher was present in the booth throughout the recordings.

3.4. Measurements

The five middle repetitions were used for measurements. The .wav files were extracted and labelled manually using the Praat software according to the following criteria and based on information from both waveforms and spectrograms:

1. Closure duration: this was defined as the period of absence of energy in the spectrogram from the end of the nasal to the burst (or if no burst appeared, the start of high frequency noise for [s]). As closures of less than 5ms can be assumed to be accidental, they were not included in the statistical analysis to avoid artificially lowering mean closure duration.

2. Burst duration of stop [t]: based on the spectrogram, the burst was measured from start of the energy band at the end of the closure or nasal if there was no closure (although bursts without closures were rare) to the high frequency noise of the fricative. If there were multiple bursts, burst duration was calculated from the beginning of the first one. Bursts of less than 1.5ms were assumed to be artifacts of measuring using Praat and were not included in the statistical analysis.
3. Vowel duration: For 'mince' and 'mints', the vowel was defined as the area of strong formants on the spectrogram between the lower energy of the word-initial and post-vocalic nasal formants. For 'prince','prints', 'quince' and 'quints', the boundary between the approximant (/r/ and /w/ respectively) and the vowel proved to be difficult to determine. For this reason, the approximant and the vowel were measured together as the area of strong formants in the spectrogram between the closure of the word-initial stop (including the burst if there was one) to the start of the weak formants of the post-vocalic nasal.

4. Post-vocalic nasal duration: the nasal was measured from the end of the strong formants of the vowel to the beginning of the burst. If the demarcation between the vowel and the nasal was unclear in the spectrogram and there were two envelopes in waveform, the nasal was measured from the beginning of the second envelope. If there was only one envelope, the nasal was measured from the start of a simplification of the waveform. If there was no burst, it was measured to the beginning of the closure and in the absence of both closure and burst, it was measured to the beginning of high-frequency noise of the fricative.

5. Length of [s]: this was defined as the area of high frequency noise in the spectrogram following the closure or burst to the end of the word.

6. Total word duration: the word was measured from the start of the first consonant to the end of the high frequency noise of the fricative.

7. The closure-to-word-length ratio (closure duration / total word duration) was calculated based on the above measurements to provide a means of verifying that a possible lengthening of the [t] closure was not an artifact of the general lengthening of the word.
3.5. Statistical Analysis

One speaker proved to have no closures of any type in the Natural condition. To avoid obtaining results dependent on that single speaker, that data was not included in the statistical analysis. The five productions of each of the remaining 11 speakers were averaged for repeated measures analyses of variance (ANOVAs). For each minimal pair (prince-prints, mince-mints, quince-quints), an ANOVA was conducted for each dependent variable: burst duration, closure duration, vowel duration, fricative duration and nasal duration and closure-to-word-length ratio. The independent variables were the nature of the [t] (Underlying, Epenthetic) and the speech style (Natural, List, Contrast). The threshold for statistical significance was set to $p<0.05$. Closures of less than 5ms and

![Figure 1. Example of segmentation for a production of 'quints' in the Contrast speech style.](image)
bursts of less than 1.5ms were not used in the ANOVAs to avoid artificially lowering mean measurements, but as the presence or absence of bursts and closures may be conditioned by the independent variables (Fourakis and Port, 1986; Arvaniti, in sub.), $\chi^2$ tests were performed to evaluate the significance of the number of times that bursts and closures did not occur as a function of each of the independent variables.

4. Results

4.1. Closure duration

For each of the word pairs (prince-prints (P), mince-mints (M), quince-quints (Q)), closure was longer for words with underlying [t] than with epenthetic [t] (P: F(1,10)=17.33, p<0.002; M: F(1,10)=33.54, p<0.001; Q: F(1,10)=13.08, p<0.005).

![Figure 2](image-url)  
**Figure 2.** Mean duration of closure and standard error for the three word pairs (prince-prints, mince-mints, quince-quints) as a function of type of [t].
Speech style was also significant for all three minimal pairs (P: F(1,10)=14.82, p<0.001; M: F(1,10)=16.53, p<0.001; Q: F(1,10)=16.2, p<0.001). Planned comparisons showed that the closure duration in the Contrast condition was longer than the Natural condition (P: F(1,10)=16.65; p<0.003; M: F(1,10)=29.12, p<0.001; Q: F(1,10)=24.94, p<0.001) and that the closure duration was longer in the Contrast condition than the List condition (P: F(1,10)=15.7, p<0.003; M: F(1,10)=14.56, p<0.004; Q: F(1,10)=16.45, p<0.003). There was no statistical difference in closure duration between the List condition and the Natural condition (P: F(1,10)=0.037; p<0.85; M: F(1,10)=0.528, p<0.48; Q:F(1,10)=0.26, p<0.87).

Figure 3. Mean duration of closure for the three word pairs (prince-prints, mince-mints, quince-quints) as a function of speech condition.
The interaction of Word (Epenthetic vs Underlying) and Style (Natural, List, Contrast) was also significant for all three minimal pairs (P: $F(2,20)=6.78$, $p<0.006$; M: $F(2,20)=12.02$, $p<0.001$; Q: $F(2,20)=13.25$, $p<0.001$).

![Graphs showing average closure duration for each word pair as a function of the Word * Style interaction](image)

**Figure 4.** Average closure duration for each word pair as a function of the Word * Style interaction

To clarify the components of the interaction, a post-hoc Tukey test was conducted for each minimal pair. The statistically significant results are detailed in Tables 1 and 2. However, for all three of the word pairs, the results were similar:

- the difference in closure duration between epenthetic and underlying [t] was statistically significant only in the Contrast condition; there was no difference between epenthetic and underlying [t] in either the Natural or List conditions.
- for both epenthetic and underlying [t] words, there was no statistically significant
difference in closure duration between the Natural and List conditions.

- closure duration was significantly longer in the Contrast condition than in the Natural and List conditions.

**Table 1.** Comparison of epenthetic and underlying [t] for closure duration; statistically significant results are shaded.

<table>
<thead>
<tr>
<th>word pair</th>
<th>Natural</th>
<th>List</th>
<th>Contrast</th>
</tr>
</thead>
<tbody>
<tr>
<td>prince-prints</td>
<td>p&lt;0.64</td>
<td>p&lt;0.88</td>
<td>p&lt;0.001</td>
</tr>
<tr>
<td>mince-mints</td>
<td>p&lt;0.99</td>
<td>p&lt;0.66</td>
<td>p&lt;0.001</td>
</tr>
<tr>
<td>quince-quints</td>
<td>p&lt;0.95</td>
<td>p&lt;0.79</td>
<td>p&lt;0.001</td>
</tr>
</tbody>
</table>

**Table 2.** Comparison of speech Style for closure duration; statistically significant results are shaded.

<table>
<thead>
<tr>
<th>word</th>
<th>Natural vs List</th>
<th>Natural vs Contrast</th>
<th>List vs Contrast</th>
</tr>
</thead>
<tbody>
<tr>
<td>prince</td>
<td>p&lt;0.99</td>
<td>p&lt;0.001</td>
<td>p&lt;0.001</td>
</tr>
<tr>
<td>prints</td>
<td>p&lt;1</td>
<td>p&lt;0.001</td>
<td>p&lt;0.001</td>
</tr>
<tr>
<td>mince</td>
<td>p&lt;0.99</td>
<td>p&lt;0.004</td>
<td>p&lt;0.003</td>
</tr>
<tr>
<td>mints</td>
<td>p&lt;0.54</td>
<td>p&lt;0.001</td>
<td>p&lt;0.001</td>
</tr>
<tr>
<td>quince</td>
<td>p&lt;0.95</td>
<td>p&lt;0.02</td>
<td>p&lt;0.002</td>
</tr>
<tr>
<td>quint</td>
<td>p&lt;0.79</td>
<td>p&lt;0.001</td>
<td>p&lt;0.001</td>
</tr>
</tbody>
</table>

These results show that although the main effect of Word showed a statistically significant difference in closure duration between words with epenthetic and underlying [t], that effect is driven entirely by the Contrast condition.
4.2. Ratio of closure duration to word length

The purpose of exploring the closure-to-word-length ratio was to verify that the apparent lengthening of the closure in the Contrast condition was not an artifact of the general lengthening of segments in careful speech. The results did show that closures were proportionately longer in the Contrast condition but they also showed that for all three word pairs, underlying [t] was longer than epenthetic [t], and in the case of the prince-prints and quince-quints pairs, there was no Word*Style interaction (and therefore the effect of the underlying form of [t] was not driven by the Contrast condition).

The ratio was larger for words with underlying [t] than epenthetic [t] (P: F(1,10)=8.33, p<0.016; M: F(1,10)=15.32, p<0.003; Q: F(1,10)=21.96, p<0.001).

Table 3. Mean of ratio and standard error as a function of type of [t]

<table>
<thead>
<tr>
<th></th>
<th>Mean</th>
<th>Standard error</th>
</tr>
</thead>
<tbody>
<tr>
<td>prince</td>
<td>0.082</td>
<td>0.009</td>
</tr>
<tr>
<td>prints</td>
<td>0.099</td>
<td>0.009</td>
</tr>
<tr>
<td>mince</td>
<td>0.088</td>
<td>0.01</td>
</tr>
<tr>
<td>mints</td>
<td>0.11</td>
<td>0.01</td>
</tr>
<tr>
<td>quince</td>
<td>0.08</td>
<td>0.01</td>
</tr>
<tr>
<td>quint</td>
<td>0.09</td>
<td>0.009</td>
</tr>
</tbody>
</table>

Speech style also affected the ratio : (P: F(1,10)=3.91, p<0.037, M: F(1,10)=4.91, p<0.018, Q: F(1,11)=4.10, p<0.032). Planned comparisons showed that the ratio was greater in the Contrast
condition than the Natural condition for mince-mints and quince-quints only (P: F(1,10)=3.91; p=0.076; M: F(1,10)=8.75, p<0.02; Q: F(1,10)=14.88, p<0.004) and that the closure duration was longer in the Contrast condition than the List condition for prince-prints and mince-mints but not quince-quints (P: F(1,10)=5.4, p<0.05; M: F(1,10)=5.8, p<0.04; Q: F(1,10)=4.25, p=0.06). The difference between the Natural condition and the List condition was not statistically significant (P: F(1,10)=0.507, p<0.49; M: F(1,10)=0.318, p<0.59; Q: F(1,10)=0.002, p=0.97).

### Table 4. Mean of ratio and standard error as a function of speech style

<table>
<thead>
<tr>
<th></th>
<th>Natural</th>
<th>List</th>
<th>Contrast</th>
</tr>
</thead>
<tbody>
<tr>
<td>prince-prints</td>
<td>Mean: 0.080</td>
<td>Mean: 0.081</td>
<td>Mean: 0.108</td>
</tr>
<tr>
<td></td>
<td>Std.err: 0.010</td>
<td>Std.err: 0.009</td>
<td>Std.err: 0.014</td>
</tr>
<tr>
<td>mince-mints</td>
<td>Mean: 0.085</td>
<td>Mean: 0.092</td>
<td>Mean: 0.120</td>
</tr>
<tr>
<td></td>
<td>Std.err: 0.018</td>
<td>Std.err: 0.007</td>
<td>Std.err: 0.016</td>
</tr>
<tr>
<td>quince-quints</td>
<td>Mean: 0.080</td>
<td>Mean: 0.082</td>
<td>Mean: 0.108</td>
</tr>
<tr>
<td></td>
<td>Std.err: 0.010</td>
<td>Std.err: 0.009</td>
<td>Std.err: 0.014</td>
</tr>
</tbody>
</table>

Only the mince-mints pair showed an interaction of Word*Style (F(2,20)=4.168, p<0.031). Underlying [t] in the Contrast condition was proportionately longer than Underlying [t] in the Natural and List conditions (p<0.001 in both cases).

### 4.3. Duration of burst

There was no statistically significant difference between bursts as a function of the underlying form of [t] or as a function of speech Style, though in two cases, the results were close to significance.
Table 5. Results of ANOVAs for Burst duration.

<table>
<thead>
<tr>
<th></th>
<th>prince-prints</th>
<th>mince-mints</th>
<th>quince-quints</th>
</tr>
</thead>
<tbody>
<tr>
<td>Word: ep [t] vs UR [t]</td>
<td>F(1,10)=1.289, p&lt;0.283</td>
<td>F(1,10)=4.649, p&lt;0.056</td>
<td>F(1,10)=.333, p&lt;0.577</td>
</tr>
<tr>
<td>Style: Nat-List-Con</td>
<td>F(1,10)=.588, p&lt;0.565</td>
<td>F(1,10)=1.142, p&lt;0.339</td>
<td>F(1,10)=2.719, p&lt;0.09</td>
</tr>
<tr>
<td>Interaction: Word*Style</td>
<td>F(1,10)=.965, p&lt;0.398</td>
<td>F(1,10)=1.032, p&lt;0.375</td>
<td>F(1,10)=2.031, p&lt;0.157</td>
</tr>
</tbody>
</table>

4.4 Occurrence of [t]

Chi-square tests showed that closures but not bursts were less likely in words with epenthetic [t]. For the effect of Style, closures were less likely in the Contrast condition than in the Naural or List conditions and less likely in the List condition than in the Natural condition. Bursts were less likely in the List condition than the Contrast condition. Bursts were also less likely in the Natural condition than in the List condition. As a function of word frequency and familiarity, the only significant difference was for closure; bursts were less likely in quin[t]s than in min[t]s.
Table 6. chi-square and p values for burst; statistically significant results are shaded.

<table>
<thead>
<tr>
<th></th>
<th>( \chi^2 )</th>
<th>p value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ep [t] – UR [t]</td>
<td>0.18</td>
<td>p&lt;1</td>
</tr>
<tr>
<td>Natural-List</td>
<td>0.08</td>
<td>p&lt;1</td>
</tr>
<tr>
<td>List-Contrast</td>
<td>8.9</td>
<td>p&lt;0.01</td>
</tr>
<tr>
<td>Natural-Contrast</td>
<td>7.32</td>
<td>p&lt;0.01</td>
</tr>
<tr>
<td>prin[t]s-min[t]s</td>
<td>0.06</td>
<td>p&lt;1</td>
</tr>
<tr>
<td>min[t]s-quin[t]s</td>
<td>6.78</td>
<td>p&lt;0.01</td>
</tr>
<tr>
<td>prin[t]s-quin[t]s</td>
<td>1.69</td>
<td>p&lt;0.2</td>
</tr>
</tbody>
</table>

Table 7. chi-square and p values for closure; statistically significant results are shaded.

<table>
<thead>
<tr>
<th></th>
<th>( \chi^2 )</th>
<th>p value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ep [t] - UR [t]</td>
<td>4.57</td>
<td>p&lt;0.05</td>
</tr>
<tr>
<td>Natural-List</td>
<td>29.8</td>
<td>p&lt;0.001</td>
</tr>
<tr>
<td>List-Contrast</td>
<td>21.01</td>
<td>p&lt;0.001</td>
</tr>
<tr>
<td>Natural-Contrast</td>
<td>88.88</td>
<td>p&lt;0.001</td>
</tr>
<tr>
<td>prin[t]s-min[t]</td>
<td>0.021</td>
<td>p&lt;1</td>
</tr>
<tr>
<td>min[t]s-quin[t]s</td>
<td>0.37</td>
<td>p&lt;1</td>
</tr>
<tr>
<td>prin[t]s-quin[t]s</td>
<td>1.13</td>
<td>p&lt;1</td>
</tr>
</tbody>
</table>
Figure 5. Occurrence of closure (and non-closure) and burst (and absence of burst) as a function of the underlying form of [t]

Figure 6. Occurrence of closure (and non-closure) and burst (and absence of burst) as a function of speech Style
Figure 7. Occurrence of closure (and non-closure) and burst (and absence of burst) as a function of the word (representing frequency and familiarity).

4.5. Nasal duration

In only one case was there a statistically significant effect on nasal duration: for the mince-mints pair, the ANOVA showed an effect of speech style. The general results are reported in Table 4. For the mince-mints pair, planned comparisons showed that the nasals in the Natural condition were significantly shorter than those in the Contrast condition (F(1,10)=6.926, p<0.025).
Table 8. Results of ANOVAs for Nasal duration; statistically significant results are shaded.

<table>
<thead>
<tr>
<th></th>
<th>prince-prints</th>
<th>mince-mints</th>
<th>quince-quints</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Word</strong></td>
<td>F(1,10)=1.793, p&lt;0.210</td>
<td>F(1,10)=2.892, p&lt;0.120</td>
<td>F(1,10)=.043, p&lt;0.839</td>
</tr>
<tr>
<td><strong>Style: Nat-List-Con</strong></td>
<td>F(1,10)=.981, p&lt;0.392</td>
<td>F(1,10)=4.291, p&lt;0.028</td>
<td>F(1,10)=.746, p&lt;0.487</td>
</tr>
<tr>
<td>*<em>Interaction:[t]<em>Style</em></em></td>
<td>F(1,10)=.899, p&lt;0.423</td>
<td>F(1,10)=1.583, p&lt;0.230</td>
<td>F(1,10)=392, p&lt;0.681</td>
</tr>
</tbody>
</table>

4.6 Vowel duration

For vowel duration, there was an effect of style for the prince-prints pair and the quince-quints pair (P: F(1,10)=4.102; p<0.032; Q: F(1,10)=6.156, p<0.008). Planned comparisons showed that vowel duration in the Contrast condition was longer than in the Natural condition (P: F(1,10)=7.497, p<0.03; Q: F(1,10)=6.511, p<0.03), and that vowel duration was longer in the Contrast condition than the List condition (P: F(1,10)=7.595, p<0.03; Q: F(1,10)=9.459, p<0.02). There was an interaction Word*Style for the mince-mints pair (F(1,10)=5.537; p<0.012). The vowel in *mince* was longer in the Contrast condition than in the Natural condition (p<0.02) or List Condition (p<0.001). The vowel in *mints* was longer in the List condition than in the Natural condition (p<0.011).
Table 9. Results of ANOVAs for Vowel duration; statistically significant results are shaded.

<table>
<thead>
<tr>
<th></th>
<th>prince-prints</th>
<th>mince-mints</th>
<th>quince-quints</th>
</tr>
</thead>
<tbody>
<tr>
<td>Word</td>
<td>F(1,10)=1.446 p&lt;0.257</td>
<td>F(1,10)=.152 p&lt;0.704</td>
<td>F(1,10)=.053 p&lt;0.822</td>
</tr>
<tr>
<td>Style: Nat-List-Con</td>
<td>F(1,10)=4.102, p&lt;0.32</td>
<td>F(1,10)=3.051 p&lt;0.070</td>
<td>F(1,10)=6.156 p&lt;0.008</td>
</tr>
<tr>
<td>Interaction: Word*Style</td>
<td>F(1,10)=3.202 p&lt;0.62</td>
<td>F(1,10)=5.537 p&lt;0.012</td>
<td>F(1,10)=.427 p&lt;0.658</td>
</tr>
</tbody>
</table>

4.7. Fricative duration

For each of the three word pairs, there is an effect of Style. Planned comparisons showed that the fricative duration in the Contrast condition was longer than in the Natural condition (P: F(1,10)=16.65; p<0.003; M: F(1,10)=29.12, p<0.001; Q: F(1,10)=24.94, p<0.001), and that fricative duration was longer in the Contrast condition than the List condition. (P: F(1,10)=15.7, p<0.003; M: F(1,10)=14.56, p<0.004; Q: F(1,10)=16.45, p<0.003). There was no statistical difference in fricative duration between the List condition and the Natural condition.
Table 10. Results of ANOVAs for fricative duration; statistically significant results are shaded.

<table>
<thead>
<tr>
<th></th>
<th>prince-prints</th>
<th>mince-mints</th>
<th>quince-quints</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Word</strong></td>
<td>F(1,10)=.05, p&lt;0.820</td>
<td>F(1,10)=1.82, p&lt;0.207</td>
<td>F(1,10)=.05, p&lt;0.831</td>
</tr>
<tr>
<td><strong>Style: Nat-List-Con</strong></td>
<td>F(1,10)=58.34, p&lt;0.001</td>
<td>F(1,10)=54.13, p&lt;0.001</td>
<td>F(1,10)=70.10, p&lt;0.001</td>
</tr>
<tr>
<td><strong>Interaction:Word*Style</strong></td>
<td>F(1,10)=1.35 p&lt;0.062</td>
<td>F(1,10)=5.33, p&lt;0.014</td>
<td>F(1,10)=.03 p&lt;0.970</td>
</tr>
</tbody>
</table>

Table 11. Mean of fricative duration and standard error for speech Style

<table>
<thead>
<tr>
<th></th>
<th>Natural</th>
<th>List</th>
<th>Contrast</th>
</tr>
</thead>
<tbody>
<tr>
<td>prince-prints</td>
<td>Mean: 80.97</td>
<td>Mean: 102.807</td>
<td>Mean: 205.225</td>
</tr>
<tr>
<td></td>
<td>Std.err: 6.423</td>
<td>Std.err: 9.874</td>
<td>Std.err: 13.924</td>
</tr>
<tr>
<td>mince-mints</td>
<td>Mean: 97.605</td>
<td>Mean: 103.177</td>
<td>Mean: 211.546</td>
</tr>
<tr>
<td></td>
<td>Std.err: 6.530</td>
<td>Std.err: 9.904</td>
<td>Std.err: 15.887</td>
</tr>
<tr>
<td>quince-quints</td>
<td>Mean: 85.303</td>
<td>Mean: 100.067</td>
<td>Mean: 207.261</td>
</tr>
<tr>
<td></td>
<td>Std.err: 7.036</td>
<td>Std.err: 8.719</td>
<td>Std.err: 13.555</td>
</tr>
</tbody>
</table>

For the *mince-mints* pair, there was a statistically significant interaction of Word*Style on fricative duration. The fricative in words with epenthetic [t] was longer in the Contrast condition than in the Natural or List conditions. The fricative in words with underlying [t] was also longer in the Contrast condition than in the Natural or List conditions. There was no statistically significant difference in fricative duration between the Natural and List conditions for words with either...
epenthetic or underlying [t].

5. Discussion

The purpose of this experiment was to explore the level of neutralization between epenthetic and underlying [t] in the context of word-final /ns/ clusters. Previous research has shown that in this context, /ns/ clusters in word final position, the neutralization is near-complete in normal speech style (Arvaniti, in sub). Therefore, the particularity of this experiment was to determine whether speakers could reliably produce a difference between underlying and epenthetic [t] when they needed to disambiguate members of a minimal pair. If the neutralization is incomplete but nearing completion, we would expect the following:

- In the Natural and List styles, there would be durational differences between underlying and epenthetic [t] only in infrequent and unfamiliar words.

- In the Contrast style, there would be a difference in duration between epenthetic and underlying [t], independently of word frequency.

- Epenthetic [t] would occur more often in frequent and familiar words in all conditions.

The predictions for frequency and familiarity are based on the findings of Arvaniti (in sub.). She
found that epenthesis was more likely in familiar and frequent words and that the difference between epenthetic and underlying [t] appeared only in unfamiliar and infrequent words, and suggested that such a pattern is indicative known patterns of sound change over time.

The most salient results concerned the [t] closure and closure-to-word-length ratio. The results showed that speakers produced longer closures for underlying [t] than epenthetic [t] but only in the Contrast condition. The closure for epenthetic [t]s in the Contrast condition was longer than for epenthetic [t]s in other conditions. The closure for underlying [t] in the Contrast condition was longer than for the underlying [t]s in other conditions. Although the results for closure duration indicated that any differences in closure duration were driven by the Contrast condition, the ratio, which is not affected by modifications in the overall length of the word, showed that the underlying [t] is proportionately longer than epenthetic [t] in all conditions and for all words.

Arvaniti (in sub.) found that closure was more likely to occur with underlying [t] than epenthetic [t] and this was also the case for the current study. However, other results for closure occurrence and burst were not consistent across the two studies. In the current study, the duration of the burst was not modified by the underlying form of the [t] as it was in the Arvaniti study. Furthermore, the prediction, based on the results of that study, that closures would be more likely and bursts less likely in familiar words was verified only for bursts; the current study showed that bursts were least likely in the *quince-quints* pair. However, the Arvaniti study included both word-medial and word-final /ns/-/nts/ clusters. As epenthesis is rarer word-medially than word-finally, her results would have been most representative of the occurrence of [t] in word-medial contexts.

No other study has examined the occurrence of epenthetic [t] as a function of speech style. Overall, both bursts and closures were more likely in careful speech as would be expected. Both
burst and closure were most likely in the Contrast condition and least likely in the Natural condition (although results for bursts did not always attain significance).

In the Fourakis and Port (1986) study, monosyllabic words with word-final /ns/-/nts/ clusters (similar to the test words used in the current study) had shorter nasals preceding underlying stops. The underlying nature of the stop did not affect the duration of the vowel or of the fricative. For the current study, neither nasals nor vowels were shorter before underlying [t]s than before epenthetic [t]s. The fricative [s] was lengthened but only in the Contrast condition and for both underlying and epenthetic stops. Recent studies have not reported shortening of nasals as a function of the underlying form of [t], so it is likely that this disparity with the Port and Fourakis study is caused by the tendency over time towards complete neutralization suggested by Arvaniti (in sub).

A possible analysis for the difference in closure duration between epenthetic [t] and underlying [t] that was observed in this study is that it was a reflection of orthographic or even morphological differences in the minimal pairs that were used as test words. The pairs of words examined in this experiment did not contrast only in the underlying form of [t]; there was also a regular spelling and morphological contrast. All the words where the [t] was assumed to be underlying had [t] in the orthography and a morpheme boundary between the nasal and the fricative: 'print + s', 'mint + s', 'quint + s'. The forms where no underlying [t] was assumed were mono-morphemic and there was no [t] in the orthography: 'mince', 'prince', 'quince'.

The fact that the fricative [s] was significantly longer in the Contrast condition (for both underlying and epenthetic [t] words) than in the Natural and List conditions could indicate that speakers were paying attention to the morpheme boundary and spelling contrast area of the word.
However, there is a lengthening of the vowel in the Contrast condition and this suggests that speakers are lengthening the word as a function of careful speech.

We do know from research on other well-attested processes of incomplete neutralization that orthography can play a role; for example, Fourakis and Iverson (1984) showed for German that the neutralization of final devoicing was complete when speakers did not see a written form of the word.

The only production study that has specifically explored the role of morpheme boundaries in incomplete neutralization is Warner et al. (2006). This study of a geminate/singleton distinction in Dutch showed that spelling rather than underlying morphological differences causes incomplete neutralization. It seems therefore that the orthography is a more likely cause of the incomplete neutralization but morphology cannot be ruled.

The difference between closure durations that appeared in the Contrast condition may also be a case of incomplete neutralization caused by the presence or absence of a [t] in the underlying form. The working hypothesis in conducting this experiment was that speakers would produce a difference in closure duration between epenthetic and underlying [t] in the Contrast condition and this prediction was fulfilled. However, there was another interesting result: the epenthetic [t]s that were produced in the Contrast condition were longer rather than shorter than those the speakers produced in the Natural and List conditions. It does not appear that in the Contrast condition, speakers were able to make a difference between epenthetic and underlying [t] by shortening or eliding the epenthetic [t] compared to other speech styles.

A possible explanation for this, and for the large difference in closure duration between epenthetic and underlying [t] in the Contrast condition, could be that speakers were striving to
accentuate the small differences that they have in other the other speech conditions. In a sense, speakers could have been doing the same thing in the Contrast condition as they were doing in the Natural and List conditions, but in an exaggerated way. This is confirmed by the fact that the ratio shows that speakers were producing proportionately longer closures for underlying [t] in both the Natural and List conditions.

However, although this experiment showed that speakers made a difference in production between underlying and epenthetic [t] in word-final /ns/ clusters in all speech styles, this still does not seem to be as clear a case for incomplete neutralization as final devoicing. Final devoicing is a much more robust effect; it seems to be stable over time, (Port and O'Dell, 1985; Port and Crawford, 1989, Warner, 2004) and the differences between underlyingly voiced and voiceless consonants is reflected not only in the realization of voicing itself, but also in vowel length and burst (Port and O'Dell, 1985; Port and Crawford, 1989). For stop epenthesis, the original Fourakis and Port (1986) study showed that nasals preceding epenthetic stops were longer than nasals preceding underlying stops and argued that this was evidence that underlying [t] was planned but epenthetic [t] was not. However, no other study for [t] epenthesis has reliably found such coarticulatory cues to underlying form.

Final devoicing has also been shown at least in some cases to persist even in the absence of orthographic cues (Dinnsen and Charles-Luce, 1984). Most importantly, the sub-phonemic cues that characterize final devoicing can be used by listeners to disambiguate, with better than chance accuracy, forms that are assumed to be surface identical (Port and O'Dell, 1985, Port and Crawford, 1989, Warner et al. 2004 and others). Ernestus and Baayen (2003) showed that speakers could even use the sub-phonemic cues produced in non-words to guess which past-tense allomorph to use. No
perception study has shown reliable evidence that listeners can use the relative differences in closure duration produced in [t] epenthesis (Warner, 2001; Shinya, 2005; Shosted, Kilpatrick and Arvaniti, in sub.).

Stop epenthesis in /ns/ clusters needs to be situated within the broader context of research on what has been called 'suspended contrasts' (Yu, 2007): unexpected differences in the realization of forms that are perceived to be surface identical.

There is evidence for at least two types of suspended contrast; near-mergers and incomplete neutralization. Near-mergers are the expression of a difference in production between lexical items that are perceived as homophonous and incomplete neutralization is a difference in the realization of underlyingly contrastive phonemes (Yu, 2007). In the case of incomplete-neutralization, the sub-phonemic cues that distinguish underlyingly contrastive forms can be used in linguistic tasks such as word disambiguation or guessing the correct form of a past-tense allomorph. This is clearly the case for final-devoicing as described above.

Near-mergers are rather more complex. Although speakers reliably produce a difference between two lexical items, the sub-phonemic cues that distinguish the two forms do not seem to be linguistically relevant to listeners.

The sub-phonemic cues produced in near-mergers have been shown to be large enough to be perceptible to listeners. This was demonstrated for the near-merger of palatalized consonants followed by a palatal glide and palatalized consonants followed by a high vowel and a palatal glide in Russian. Diehm and Johnson (1997), showed that native speakers produced a difference between these two sequences but could not hear that difference. This was particularly surprising because non-native speakers could hear the difference between the two forms. Diehm and Johnson proposed
that this difference was caused by a difference in attention to linguistic status; the native speakers
treated the difference as linguistically unimportant.

In a study of vowels in two dialects of Swedish, one where the vowels of 'sett' and 'satt' were
merged and one where they were not, Janson and Schulman (1983) showed that although speakers
of the dialect in which the sounds were not completely merged could produce a difference between
'satt' and 'sett', they were not able to perceive that difference. They were, however, able to perceive
that same difference in English words. Janson and Schulman suggested that this was because the
Stockholm dialect where the vowels were merged was dominant and that speakers of the minority
dialect had become used to ignoring the sub-phonemic cues as linguistically irrelevant for Swedish.

These studies indicate that the sub-phonemic cues produced in near-mergers do not seem to
encode linguistic information and, in the case of the Janson and Schulman study, that geographic
dialect may play a role in maintaining the merger. There is evidence that other factors such as
orthography, morpho-syntact and social dialect may play a role in near-mergers.

As discussed previously, orthography has long been a suspect in maintaining suspended
contrasts. In situations where the difference in underlying form is represented in the orthography,
methodologies that minimize use of written materials consistently find smaller effects than those
that use written prompts (Charles-Luce, 1985; Fourakis and Iverson, 1989).

Yu (2007) suggested that morpho-syntact also can play a role in maintaining near-mergers.
Yu showed that in Cantonese, where phonology is not reflected in orthography, lexical mid-rising
tone was realized differently from morphologically derived mid-rising tone. In this case, the
difference appears to be maintained at the morphological level.

Finally, Hay et al. (2006) investigated the perception of the evolving near-merger of 'NEAR'
and 'SQUARE' diphthongs in New Zealand English. Again, the difference between the diphthongs is not regularly reflected in the orthography. Hay et al. showed that listeners performed differently as a function of the social context with which the stimuli were associated. They argued that listeners are sensitive to information about social class in the speech signal. It is not clear for the 'Near Square' merger in New Zealand English whether the differences in production between the two diphthongs will become completely merged over time or if they will become more markers of a social dialect.

Some researchers believe that the disjunction that is observed between production and perception in near-mergers can be best explained within the framework of exemplar theory (Pierrehumbert, 2002; Hay et al, 2006; Yu, 2007 and others). In an exemplar model, instances of a phoneme are stored as a cloud. These exemplars would include both fine phonetic detail and dialectal information. As sounds change and converge, the clouds start to overlap. In production, the target is an averaging of exemplars within for a given phoneme. As long as the two clouds do not completely overlap, the productions for those two sounds are predicted to be slightly different. However, for perception, without contextual and social information, near-merged tokens will fall within the overlapping area of the two clouds.

Situating [t] epenthesis in /ns/ clusters within the body of research on incomplete neutralization and near-mergers is not straight-forward. The apparent non-availability for linguistic purposes of the sub-phonemic cues that distinguish underlying and epenthetic [t] seems to favor near-merger. However, [t] epenthesis is usually modeled as a phonological process and this is assumed to be a characteristic of incomplete neutralization (Yu, 2007). There are some important differences between the phonological rules of [t] epenthesis in /ns/ clusters and final-devoicing. For
final-devoicing, the realization of segments alternates as a function of the environment.

(6) Example of final devoicing alternation for Dutch (from Ernestus and Baayen, 2003)

<table>
<thead>
<tr>
<th></th>
<th>Dutch Word</th>
<th>IPA</th>
<th>Function</th>
</tr>
</thead>
<tbody>
<tr>
<td>a.</td>
<td>verwijden</td>
<td>[vɛrˈvɪdə(n)]</td>
<td>'widen-inf'</td>
</tr>
<tr>
<td>b.</td>
<td>verwijtec</td>
<td>[vɛrˈvɪtə(n)]</td>
<td>'reproach-inf'</td>
</tr>
<tr>
<td>c.</td>
<td>verwijd</td>
<td>[vɛrˈvɪt]</td>
<td>'widen'</td>
</tr>
<tr>
<td>d.</td>
<td>verwijt</td>
<td>[vɛrˈvɪt]</td>
<td>'reproach'</td>
</tr>
</tbody>
</table>

In the case of [t] epenthesis, the only way to show whether the rule had been applied or not would be to compare it with dialects in which [t] epenthesis does not occur. Furthermore, for final devoicing, the underlying form of a segment can be determined through the alternations, whereas in [t] epenthesis, we are forced to depend on spelling to determine whether a stop is underlying or epenthetic.

There are also articulatory differences. There is an assumption that for final-devoicing, a segment that is underlyingly voiced and realized as voiceless is articulated in the same way as an underlyingly voiceless segment realized as voiceless. However, for [t] epenthesis in /ns/ clusters, there is evidence that epenthetic [t]s may be articulated differently to underlying [t]s (Ali et al. 1979).

Finally, for final devoicing the vowel preceding an underlyingly voiced consonant realized as voiceless has been shown to be longer than for an underlyingly voiceless consonant realized as voiceless (Port and O'Dell, 1985) and there are also more subtle coarticulatory effects that are not singly statistically significant (Port and Crawford, 1989). No study on [t] epenthesis in /ns/ clusters since Fourakis and Port (1986) has shown reliable evidence of measurable coarticulatory effects.
Considered within the context of suspended contrasts, [t] epenthesis in /ns/ clusters patterns more clearly with near-mergers than with incomplete neutralization. However, there is still work to be done. Future work could explore the linguistic status of the sub-phonemic cues that are produced. An 'AX' experiment would determine whether the sub-phonemic cues can have a linguistic function. In AX experiments, listeners are asked whether two words played one after the other are the same or different. If [t] epenthesis in /ns/ clusters is a near-merger, we would expect that listeners would not be able to hear the difference between words with epenthetic and underlying [t] in an AX experiment.

Determining the cause of suspended contrast for [t] epenthesis will be challenging. The three causes of near-merger that have been suggested, orthography, morpho-syntax and sociolinguistics, may all apply for [t] epenthesis. It is not possible to eliminate orthography because our assumptions about the presence or absence of underlying [t] are based on the spelling. For morphology, the minimal pairs in which [t] epenthesis occurs most reliably (word final /ns/ clusters) are morphological alternations. Investigating other environments (for words like 'fancy') is a possibility but will be handicapped by the difficulty of finding minimal pairs and the low occurrence of epenthetic [t] word-medially.

There have been no reports that epenthetic [t] in /ns/ clusters is affected by social class, but we do know that it is a geographic dialect variation (Fourakis and Port, 1986). Further research should examine the perception of epenthetic [t] in the context of social and geographic information.
6. Conclusion

The purpose of this experiment was to examine whether speakers could produce a difference between underlying and epenthetic [t] in [ns] clusters when they were forced to make a contrast between members of a minimal pair and this was shown to be the case. Furthermore, use of closure-to-word length ratio in conjunction with raw measurements of closure showed that speakers produced a difference between underlying and epenthetic [t] even in other speech styles. However, this evidence considered within the broader context of other studies, seems to indicate that [t] epenthesis may be a case of near-merger rather than incomplete-neutralization.

It is assumed that in normal language change, near-merged sounds become merged over time. However, in some cases, and this seems to be the case for [t] epenthesis in /ns/ clusters, complete merging is blocked. Orthography, morphology and dialect have all been shown to play a role in maintaining a near-mergers. Future research on [t] epenthesis in /ns/ clusters will need to determine whether the difference in production is maintained only by spelling (and possibly morphology) or if listeners can use the sub-phonemic cues to identify speakers from other dialect groups.
7. References


Arvaniti, A. Submitted. Incomplete neutralization: the case of epenthetic [t].


Shosted, R., Kilpatrick, C., Arvaniti, A. Submitted. On the perception of incomplete neutralization.


