

# Duration of epenthetic [t] in polysyllabic American English words

**Isaiah WonHo Yoo**

Department of Foreign Languages and Literatures, MIT  
*iyoo@mit.edu*

**Barbara Blankenship**

Department of Linguistics, UCLA  
*blankens@humnet.ucla.edu*

This paper reports how stress and the position of the /ns/ cluster in polysyllabic words affect [t] epenthesis in four different environments: (1) word-medial after a stressed vowel (e.g. 'cen.sus vs. 'dents us); (2) word-final after a stressed vowel (e.g. in.'tense vs. in'tents); (3) word-medial after a stressless vowel (e.g. con.'sent vs. blunt'say.ing); and (4) word-final after a stressless vowel (e.g. 'sci.ence vs. 'con.tents). Analysis of stop closure durations in experimental sentences read by seven American English speakers reveals that position, not stress, is the most important factor in [t] epenthesis: final position (e.g. *science* and *intense*) favors epenthesis. Stress is found to have an effect on stop closure durations in the way it interacted with word-position – i.e. for the final /ns/ cluster, stress immediately before it disfavors epenthesis (e.g. *intense*). The underlying /t/ is shown to be not significantly longer than the epenthetic [t]. Measurements from polysyllabic words in the TIMIT corpus corroborate the experimental results that word-final position favors epenthesis and that stress does not correlate with epenthesis. In the TIMIT data, however, underlying /t/ was significantly longer than epenthetic [t].

## 1 Introduction

This paper examines the epenthesis of a [t] in the consonant cluster environment /ns/. The insertion of a [t] between an /n/ and an /s/ – producing the surface form [prnts] from the underlying representation /prms/, for example – is a well-known phonological phenomenon in American English (Zwicky 1972, Barnitz 1974, Ohala 1974, Ali et al. 1979, Dinnsen 1980, Fourakis & Port 1986, Clements 1987, Lee 1990, Blankenship 1992). Most research on this process, however, has been restricted to examining monosyllabic words that constitute minimal pairs such as *prince* vs. *prints* and *tense* vs. *tents* (Zwicky 1972, Dinnsen 1980). What, then, happens when /ns/ occurs in a polysyllabic word? Does the epenthesis of a [t] occur as regularly as it does in monosyllabic words? Or could the process be affected by other factors such as stress and the position of the /ns/ cluster?

The examples in (1) illustrate how epenthesis neutralizes the phonemic distinction between word pairs:

- (1) a. [prnts]     *prince/prints*  
    b. [dents]     *dense/dents*

- c. [tɛnts]      *tense/tents*  
 d. [sɛnts]      *sense/cents*

Fourakis & Port (1986), however, found that the duration of an epenthetic [t] was shorter than that of an underlying /t/, thus maintaining that the supposed neutralization shown in the above examples is a case of incomplete neutralization. Fourakis & Port also found that only the American consultants in their experiment exhibited the epenthetic [t] (which they defined as an interval greater than 10 ms long with no acoustic energy visible on a spectrogram), while their South African consultants made clear distinctions between minimal pairs such as *dents* and *dense*. This finding provides evidence against such claims as Ohala's (1974) that stop epenthesis is governed by articulatory constraints universal to all languages.

## 2 The Stop epenthesis rule

Dinnsen (1980) formulates a basic rule to account for not only the epenthetic [t] but also other epenthetic stops in any nasal-fricative cluster environment, as in [lɛŋkθ] *length*, [mʌntθ] *month*, and [lɪmpf] *lymph*. Fourakis & Port (1986) modify the rule by addition of a syllable boundary:

$$\begin{array}{ccc}
 \left[ \begin{array}{c} + \text{ con} \\ + \text{ nas} \end{array} \right] & \left[ \begin{array}{c} - \text{ son} \\ + \text{ cont} \\ \alpha \text{ voi} \end{array} \right] & \$ \quad 1 \quad 1 \quad 2 \quad 3 \\
 & & \rightarrow \left[ \begin{array}{c} - \text{ nas} \\ \alpha \text{ voi} \end{array} \right] \\
 1 & 2 & 3
 \end{array}$$

In the Fourakis & Port account, syllable boundaries play an important role in determining whether an epenthetic stop occurs in surface forms; that is, epenthesis applies only when the sonorant and the fricative are in the same syllable.

Blankenship (1992), examining tokens retrieved from the TIMIT corpus, tested how stress might affect stop epenthesis in read sentences of American English. She found that epenthesis, although it could occur in any stress environment, rarely occurred when the following vowel was stressed. This finding provides evidence for Clements's (1987) account of stop epenthesis (or intrusive stops, in his terms). Formulated within the framework of an autosegmental analysis, Clements's (1987: 32) rule predicts that epenthesis does not occur 'if the following consonant is in a stressed syllable, as in *consort*, *consortium*', although it is 'possible, though probably not obligatory', before a stressless syllable, as in *cancer* or *concert*. Arguing against the use of the rule format  $\emptyset \rightarrow B / C\_D$ , Clements maintains that the durational specification for epenthetic stops can be better represented if we assume that the epenthetic [t] shares a C-slot with the following fricative /s/, while the underlying /t/ occupies its own.

These two studies demonstrate that stress and syllabification are related factors and that stress partly determines syllables. With following stress, the /s/ will be in the onset of the following syllable, resulting in no epenthesis; whereas with preceding stress, the /s/ will likely be incorporated into the coda of the previous syllable, resulting in epenthesis.

A further issue about epenthetic stops is whether they are shorter in duration than the underlying ones. Blankenship (1992) found, contrary to Fourakis & Port's (1986) result, that in TIMIT there was no significant difference in duration between the epenthetic [t] and the underlying /t/. This result lends support to Lee's (1990) earlier finding that minimal pairs such as *intense* and *in tents* exhibit no difference in their phonetic realizations. Noting that consistent differences between the underlying /t/ and the epenthetic [t] were found only in

**Table 1** /ns/ words.

	Word-medial	Word-final
After a stressed vowel	'cen.sus/'sen.si.ble 'sen.so.ry/'den.si.ty	in.'tense/in.'cense de.'fense/of.'fense
After a stressless vowel	con.'sent/con.'ceal con.'sole/con.'sist	'sci.ence/'pres.ence in.'sur.ance/'pre.science

the pair *dense/dents* and not in the other similar pairs in Fourakis & Port's study, Lee asserts that the claim of a categorical duration difference lacks adequate support from empirical findings.

Attributing Blankenship's (1992) failure to replicate Fourakis & Port's (1986) result to the variability of corpus data, Warner & Weber (2001) maintain that the epenthetic stops and the underlying stops do exhibit different durations. Their experiment consisted of 14 read monosyllabic Dutch nonsense words created by using the CELEX database. Listeners who were asked to press a button as quickly as possible when they heard a particular stop were far slower to respond to epenthetic stops than they were to underlying stops. Warner & Weber argue, in support of Fourakis & Port's finding, that the listeners perceived the epenthetic stops more slowly because they were not acoustically as strong as the underlying ones. Perceptual strength was associated with burst power and duration and with durations of the preceding nasal, the silence before the burst, and the total closure.

### 3 Hypotheses

To examine the effects of stress and syllable affiliation on epenthesis in /ns/ clusters, polysyllabic words were selected. The following hypotheses were tested:

1. Epenthetic [t] in a polysyllabic word is longer word-finally than word-medially.
2. Epenthetic [t] in a polysyllabic word is longer after a stressed vowel than after a stressless vowel.
3. The duration of an underlying /t/ is longer than that of an epenthetic [t] in polysyllabic words.

### 4 Methods

The words used in this experiment provide /ns/ clusters in four different environments: (1) word-medial after a stressed vowel; (2) word-final after a stressed vowel; (3) word-medial after a stressless vowel; and (4) word-final after a stressless vowel. The /ns/ words are compared with words with /nts/ clusters in the same environments. Because /nts/ does not occur in the middle of a morpheme, the /nts/ samples are not always single words. The word-medial samples have /nts/ medial to an adjective-noun or verb-object group instead of to an individual word. There are two /ns/ and one /nts/ words in each environment. Tables 1 and 2 show both sets of words in their respective environments.

Seven native speakers of English, four males and three females, participated in this experiment. Each speaker read the words in the Word List (listed in appendix 1) twice in the carrier sentence 'Say \_\_\_ again'. The Word List was created by randomizing all the words in tables 1 and 2. The recordings were made in a soundproof booth located in the UCLA

**Table 2** /nts/ words and phrases.

	Word-medial	Word-final
After a stressed vowel	'dents us/ 'punts us	in 'tents/non-'cents
After a stressless vowel	blunt 'say.ing/spent 'sav.ings	'con.tents/'pres.ents

phonetics lab. Subsequently, the recordings were digitized and waveforms and spectrograms of the words displayed with the PCquirer program. Voiceless closure was taken to be the sign of a stop, and the duration of any such closure was measured and compared across conditions, as was done by Fourakis & Port (1986). It was difficult to pinpoint the end of the preceding nasal /n/ in many tokens. To maintain consistency in measuring the durations, any visible remnant of voicing before the silence was considered part of the preceding nasal. The end of the stop closure was measured where any visible fricative noise was seen in the spectrograms.

Since there are three factors involved in this experiment – i.e. position, stress, and /t/ status (/ns/ or /nts/) – a three-factor repeated-measures ANOVA was conducted in order to test hypotheses 1 and 2. The analysis, however, showed neither main effect for /t/ status (/ns/ or /nts/) nor any interactions involving /t/ status. Therefore, we collapsed across /t/ status by averaging and performed a two-factor repeated-measures ANOVA to analyze the /ns/ words. To test hypothesis 3, another ANOVA was conducted, now including the /nts/ words and with three factors: position, stress, and /t/ status (/ns/ or /nts/), whose durations were averaged across both items and productions.

## 5 Results

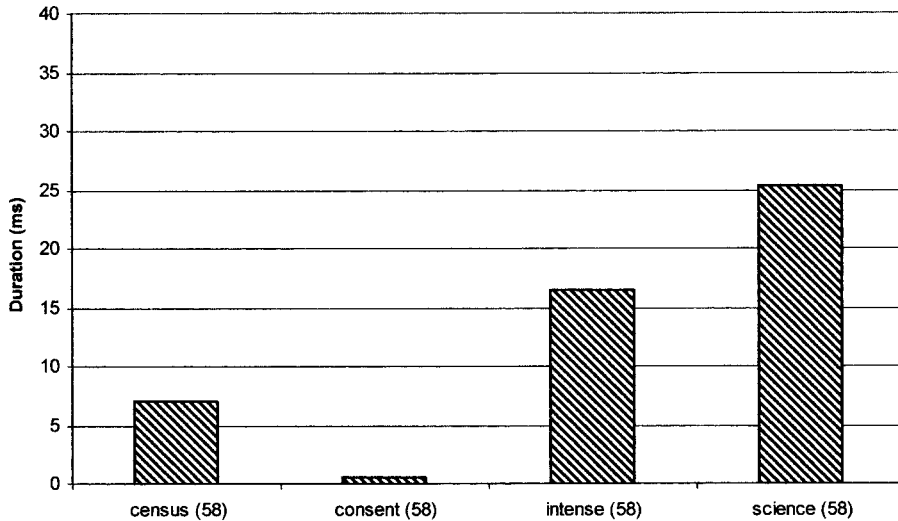
### 5.1 Hypothesis 1: medial vs. final position

Figure 1 graphs the stop closure durations of the /ns/ words in each group averaged across the seven speakers. As was predicted by hypothesis 1, the word-medial groups, i.e. *census* and *consent*, exhibited virtually no stop closure durations (7.07 ms and 0.64 ms, respectively), whereas the word-final groups, i.e. *intense* and *science*, exhibited longer durations (16.48 ms and 25.31 ms, respectively). Another difference between the two groups is that in the first pair, the *s* is syllabified with the second vowel, while in the second pair, the *s* is in the same syllable as the *n*. Thus, it seems to be the case that whenever the /ns/ cluster lies within a single syllable, a full epenthetic [t] (at least 15 ms duration) occurs, but whenever the *n* and *s* are allied with different syllables, a very short epenthetic [t] is produced.

While all of the durations are small when averaged over productions, there are reliable differences among the conditions. As shown in table 3 below, the two-factor repeated-measures ANOVA conducted on the /ns/ words, with durations being the dependent variable, revealed an overall effect of word-position (medial or final), with final position favoring epenthesis. All significance judgments are based on  $p < .05$ .

### 5.2 Hypothesis 2: stressed vs. stressless environment

Although there is no main effect of stress, the ANOVA also shows an interaction effect between the two factors: position and stress. Subsequent t-tests on the simple effect of stress for the medial clusters (i.e. *census* and *consent*) and for the final clusters (i.e. *intense* and *science*) proved to be statistically significant, both pairs exhibiting the  $p$  value less than



**Figure 1** Average of [t] closure durations in /ns/ words. Labels identify the stress and word-position categories; the number of instances is given in parentheses.

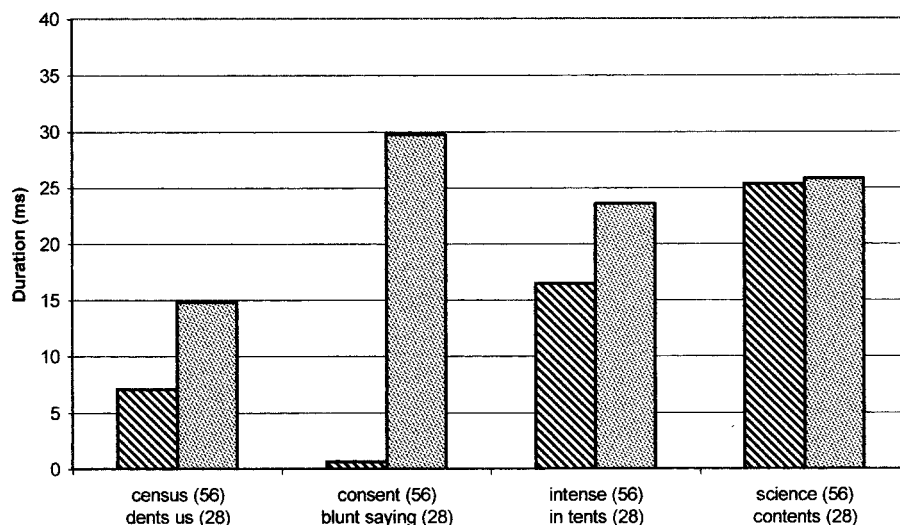
**Table 3** Hypotheses 1 and 2: two-factor ANOVA.

Source	SS	df	MS	F	Sig.
Position	8666.241	1	8666.241	55.986	.000
Stress	85.226	1	85.226	.551	.460
Position*Stress	1403.681	1	1403.681	9.068	.003
Error	16717.577	108	154.792		
Total	26872.725	111			

.025, a p value adjusted for multiple t-tests with the Bonferroni correction. The reason there is an interaction effect between position and stress, but no main effect of stress, is that stress works oppositely in the two positions: stress before the medial /ns/ cluster favors epenthesis, the *census* group exhibiting a longer stop closure duration (7.07 ms) than that of the *consent* group (0.64 ms); stress before the final /ns/ cluster disfavors epenthesis, the *intense* group exhibiting a shorter stop closure duration (16.48 ms) than that of the *science* group (25.31 ms), although both final clusters show more epenthesis than either medial cluster.

### 5.3 Hypothesis 3: duration of underlying /t/ vs. epenthetic [t]

Figure 2 graphs the stop closure durations of the /ns/ and /nts/ words in each group averaged across the seven speakers. The stop closure durations of the /nts/ items in all four environments are longer than those of the equivalent /ns/ items. But the differences in durations between the /ns/ and /nts/ words are not consistent across the four different environments: the second and the fourth groups exhibit different patterns from those of the first and the third groups and from each other. Unlike the epenthetic ones, the underlying stops are all fairly long and do not differ as much in duration depending on context, which



**Figure 2** Average of [t] closure durations in /ns/ (striped bars) vs. /nts/ (shaded bars) words. Labels identify the stress and word-position categories; the number of instances is given in parentheses.

**Table 4** Hypothesis 3: three-factor ANOVA.

Source	SS	df	MS	F	Sig.
Position	1326.004	1	1326.004	8.630	.005
Stress	334.305	1	334.305	2.176	.147
[t] status	1729.383	1	1729.383	11.256	.002
Position*Stress	5.738	1	5.738	.037	.848
Position*[t] status	748.252	1	748.252	4.870	.032
Stress*[t] status	190.828	1	190.828	1.242	.271
Position*Stress*[t] status	684.076	1	684.076	4.452	.040
Error	7374.903	48	153.644		
Total	12393.489	55			

may have stemmed from the fact that the underlying /t/s all have either morpheme or word boundaries after them, while none of the epenthetic ones do. Also worth noting is the fact that there seems to be stop-weakening in *dents us*: the average stop closure duration of the *dents us* group is only 14.8 ms, considerably shorter than those of the other three groups.

As can be seen in table 4, the three-factor repeated-measures ANOVA conducted on the /ns/ and the /nts/ words revealed overall effects of word-position and also of /t/ status (/ns/ or /nts/). As with the /ns/ words alone, there is no main effect of stress.

The ANOVA shows two interaction effects between the three factors tested: (1) between position and /t/ status and (2) between position, stress, and [t] status. Therefore, t-tests were performed to determine where the interactions were at work. Only one of the four differences within pairs (medial-stressless pair, e.g. *consent/blunt saying*) proved to be statistically significant, with a p value less than .0125 (a p value adjusted for multiple t-tests with the Bonferroni correction). Another pair (medial-stressed pair, e.g. *census/dents*

us) also exhibited a low, if not statistically significant,  $p = .067$ . But for the aforementioned stop-weakening in the *dents us* group, it would be reasonable to assume that the difference between *census* and *dents us* might have also proved statistically significant.

## 6 Summary of experimental results

Hypothesis 1 – epenthetic [t] in a polysyllabic word is longer word-finally than word-medially – is supported: the word-final groups, *intense* and *science*, had significantly longer stop closure durations than did the word-medial groups, *census* and *consent*.

Hypothesis 2 – epenthetic [t] in a polysyllabic word is longer after a stressed vowel than after a stressless vowel – is not supported. Stress by itself was found to have no effect. For the final /ns/ cluster, epenthetic [t] is significantly shorter after a stressed vowel, i.e. *intense*, than after a stressless vowel, i.e. *science*. However, for the medial /ns/ cluster, epenthetic [t] was significantly longer after a stressed vowel, i.e. *census*, than after a stressless vowel, i.e. *consent*, as predicted by the hypothesis. But neither *census* nor *consent* would likely sound epentheticized, as the epenthetic [t]s in both groups averaged less than 8 ms.

Hypothesis 3 – the duration of an underlying /t/ is longer than that of an epenthetic [t] in polysyllabic words – is supported only ‘apparently’ by the main effect of /t/ status, in which the average of the underlying /t/ (23.49 ms) was significantly longer than that of the epenthetic [t] (12.38 ms). A closer examination of each of the four different environments revealed that the difference does not hold true for every group tested in this experiment (figure 2). Only the difference in the second group – *blunt saying* vs. *consent* – was found to be statistically significant in the subsequent t-tests. This result is especially important because this is the case in which the /ns/ words have no epenthesis at all (mean < 1 ms). In the cases where there was some epenthesis, the durations of the epenthetic [t]s proved to be not significantly shorter than those of the underlying /t/s in /nts/ words. Thus, the relevant question for this experiment is whether there is any epenthesis at all. If there is, the durations between the underlying and the epenthetic stops are not significantly different. The fact that a large difference in only one group can skew the representation of all the groups emphasizes the importance of examining various environments before one makes a general conclusion about any particular phenomenon.

## 7 Corroborating data

The TIMIT corpus provides phonetic transcriptions and segment durations for sentences read by 630 native speakers of American English. Each speaker read 10 items from a set of 2342 research sentences, providing a corpus of 6300 utterances. Appendix 2 summarizes the types of research sentences and the allocation of sentences to readers. The transcriptions are consistent and accurate. They are based on computer-aided segmentation corrected by humans using both spectrograms and recordings.

TIMIT transcriptions of the segment /t/ provide separate duration measurements for the closure and the release. There are frequent instances of closures without a release and releases without a closure. The sum of the closure and release durations is used as [t] duration in the analysis that follows. Thus, a duration of 0 indicates that neither a closure nor a release was observed by the transcribers.

The TIMIT data used in Blankenship (1992) included /ns/ and /nts/ sequences within both monosyllabic and polysyllabic words, as well as over word boundaries (e.g. n#s or ns#t). To make the corpus data more comparable with our experimental results, a new set of tokens

**Table 5** Number of qualifying /ns/ and /nts/ tokens in the TIMIT corpus. The number and the percentage of tokens realized with a [t] in each category is shown in parentheses.

	Word-medial	Word-final
Words with /ns/		
After a stressed vowel	69 (20, 29%)	9 (5, 56%)
After a stressless vowel	62 (6, 10%)	80 (27, 34%)
Words with /nts/		
After a stressed vowel	7 (6, 86%)	21 (19, 90%)
After a stressless vowel	31 (29, 94%)	35 (25, 71%)

was selected from the TIMIT corpus according to the following criteria:

- The target sequence /ns/ or /nts/ was within a single polysyllabic word.
- Words with [t] immediately following the target sequence (e.g. *instant*) were excluded to eliminate a possible influence of that *t* on the epenthesis process.
- Words at the end of an utterance were eliminated because all segments are lengthened in that position.
- Words with [ns] or [nts] transcriptions resulting from underlying segments other than /n/, /t/, or /s/ (e.g. *months*) were excluded.
- Words where the underlying segments were realized with phones other than [n], [t], or [s] (e.g. *compensation* rendered [as kəmpɛnsɛʃən]) were excluded.

The selected words were aggregated by stress pattern and position within the word. Table 5 shows the number of tokens in each category and the percentage of those tokens that were realized with any kind of [t] (any non-zero value for closure or release duration). Among the /ns/ words, being either word-final or a stressed vowel favored the presence of a [t] element. When only the presence or absence of [t] is considered, the percentages in table 5 display the relationships suggested in hypotheses 1 and 2. The actual durations of the [t] elements, however, do not support the hypotheses, as will be shown below.

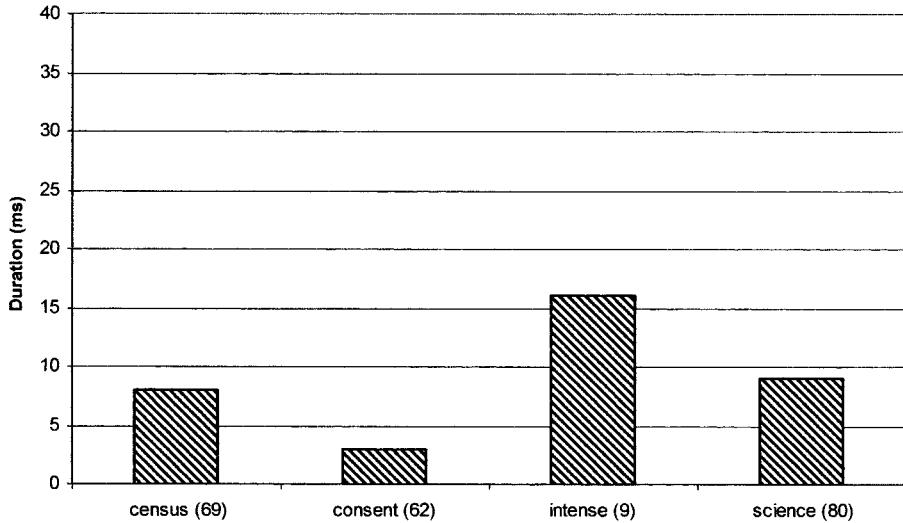
It is surprising that so many of the /nts/ words in TIMIT were pronounced without [t], given that all of the /nts/ words in our experiment were pronounced with [t]. The effect may be due to differences in the recording stimuli. The frame sentence in the experiment causes readers to be more aware of special status of the target word and perhaps pronounce it more carefully. TIMIT speakers were reading sentences of unrelated structure and thus were unaware that any particular word would be under study.

Figure 3 shows the average [t] durations in the /ns/ words in each category. Zero durations are included in the averages. For convenience, the categories bear the same labels as figure 1, although those particular words were not in the TIMIT corpus.

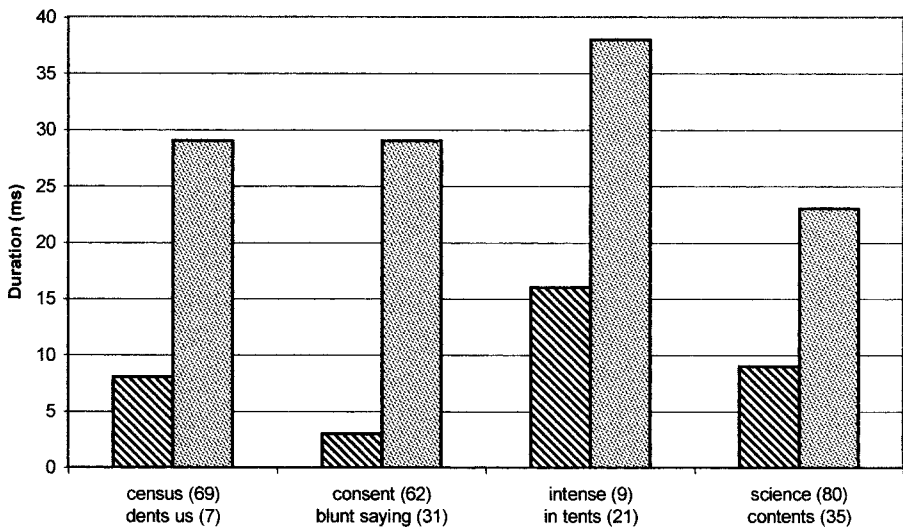
Due to unmatched data sets, an analysis of variance is not possible, but 1-tailed t-tests assuming unequal variance showed that position had a significant effect on [t] duration, supporting hypothesis 1. The effect due to stress that is apparent in the figure is not significant. Thus hypothesis 2 is not supported statistically by the TIMIT duration data. These two results corroborate our experimental data for hypotheses 1 and 2.

Figure 4 shows the average stop closure durations of the TIMIT /ns/ and /nts/ words by category. As in figure 2, average [t] durations of the /nts/ words in all four environments appear longer than those of the /ns/ words. Based on t-tests ( $p < .05$ ), the difference in average durations is significant in all four groups. Thus, hypothesis 3 is supported statistically. Similarly to the experimental data, the duration difference is largest within the pre-stress medial group (consent/blunt saying). But where the experiment found significant differences only in that group, the TIMIT data exhibit significant differences in each group.





**Figure 3** Average of [t] closure durations in TIMIT /ns/ words. Category labels match those of figure 1 and do not represent actual words in the TIMIT corpus; the number of instances is given in parentheses.



**Figure 4** Average of [t] closure durations in TIMIT /ns/ (striped bars) vs. /nts/ (shaded bars) words. Category labels match those of figure 1 and do not represent actual words in the TIMIT corpus; the number of instances is given in parentheses.

## 8 Discussion

A crucial factor in Fourakis & Port's (1986) version of Dinnsen's (1980) rule is that the nasal-fragative cluster must occur before a syllable boundary in order for an epenthetic stop to occur. The rule proves to be a good generalization for epenthetic [t] in both our experimental data and the TIMIT corpus. The words in the *intense* and *science* groups exhibited longer epenthetic [t] durations. Average durations were shorter in the *census* and *consent* groups because the /ns/

clusters in these words occur across syllable boundaries. The longer average [t] duration in the *census* group than in the *consent* group may be attributed to some speakers' resyllabifying in *census* due to the stress and thus producing an epenthetic [t] in the /ns/ cluster according to the rule. This individual difference, in fact, can be seen among the seven speakers who participated in our experiment. Unlike the *consent* group, which had only two speakers with any measurable stop durations (0.75 ms and 3.70 ms), the *census* group presented much more diverse results: one speaker produced an unambiguous epenthetic [t] (24.85 ms), another produced no epenthesis at all, and the others produced measurable stop durations under 10 ms. Our results would equally well support a theory that the *s* is ambisyllabic.

## 9 Conclusions

This paper examined whether an epenthetic [t] in an /ns/ cluster, a well-documented phenomenon in American English monosyllabic words, also occurs in polysyllabic words and whether there is any difference in durations between the underlying /t/ and the epenthetic [t] in such words. The experiment has revealed that epenthesis can occur in any stress or word position, but is likely to be longer when it is at the end of a word. This result is in line with Blankenship's (1992) finding that epenthesis could occur in any stress environment and is substantiated by our new data set from the TIMIT corpus. The degree to which these results support Fourakis & Port's (1986) version of Dinnsen's (1980) epenthesis rule was discussed.

Comparison of the /ns/ and /nts/ words has revealed that, on average, the stop closure duration of an underlying /t/ is longer than that of an epenthetic [t]. This seems to lend support to hypothesis 3, which was formulated with reference to Fourakis & Port's earlier finding. However, while our TIMIT data exhibited significant duration differences in all four environments, the experiment yielded in a significant difference in only one environment, *consent* vs. *blunt saying*. The fact that results from one of the experimental environments can skew the representation of the combined data should serve as a warning against making generalizations based on only one environment.

The results of our experiment support Lee's (1990) finding that the minimal pair *intense* and *in tents* (the third group compared with regard to hypothesis 3) exhibited no significant difference in their surface forms, but our TIMIT data support the opposing view of Fourakis & Port on this question. (Lee's experiment examined 54 /ns/ and 54 /nts/ instances and Fourakis & Port's 60 /ns/ and 60 /nts/ monosyllabic instances from their American speakers; our experiment examined 56 /ns/ and 28 /nts/ polysyllabic instances in the *intense/in tents* group; and the TIMIT set contained 9 /ns/ and 21 /nts/ polysyllabic words in that group.) One related question that this experiment did not address is the perception of the epenthetic [t]. When told after the experiment that this study was about the epenthetic [t] in the /ns/ environment, most of the participants said that they could hear an epenthetic [t] in their own pronunciation of *census*. What accounts for this discrepancy between their perception and the acoustic signals? We will elaborate on two likely answers to this question: the effect of the silence preceding the /s/ in the /ns/ cluster and individual speaker variation.

SILENCE. Given an adequate length of silence, listeners tend to perceive a stop consonant even when there is no acoustic burst. In this experiment, many words in both the /nts/ and the /ns/ sets did not exhibit stop bursts in their acoustic signals. Without stop bursts, the *census* group's average 7.07 ms of silence between the /n/ and the /s/ must have contributed to the consultants' perception of an epenthetic [t]. In their perception experiment, Warner & Weber (2001) also found that some tokens where listeners reported hearing epenthetic stops had no epenthetic bursts visible in the spectrograms. They concluded that closure duration might be a contributing factor for perception of epenthetic stops. Likewise, Ali et al. (1979) found that epenthetic stops were cued perceptually by silent gaps, with or without a burst release. Their data included a few instances of monosyllables where epenthesis was perceived in the

complete absence of a silent gap or burst, and polysyllables where silent gaps elicited no perception of epenthesis at the syllable juncture. They concluded that since listeners expect some sort of gap at junctures, they are less disposed to interpret such gaps as an epenthetic element.

**SPEAKER VARIATION.** Although Fourakis & Port's (1986) American speakers always had epenthetic [t]'s in monosyllabic words such as *prince* and *dense*, there could be individual variation for the same environment in polysyllabic words. In support of this possibility are Blankenship's (1992) observations that only about one-fourth of the /ns/ consonant clusters in the TIMIT corpus of American English exhibit epenthesis and that speakers who produced epenthesis did not exhibit it consistently. The individual difference seen in the *census* group in this experiment further supports this possibility. Other studies show that even monosyllables are subject to variation. Barnitz (1974) and Ali et al. (1979) found that not all subjects epenthesized the same words, even in monosyllables. Noting the fact that their listeners perceived epenthetic stops far less often than they perceived underlying stops, Warner & Weber (2001) also attribute this perceptual discrepancy to the possibility that their speakers might not have been consistent in producing epenthetic stops. More empirical studies examining the other nasal-fricative clusters are needed to arrive at a better understanding of how prevalent these individual variations are both in monosyllabic and in polysyllabic words.

Further research is also required on the articulatory details of epenthesis, particularly with regard to speaker variability. In the opposite articulatory process, where a consonant cluster is simplified by phoneme deletion (for example, the deletion of [t] over the word boundary in 'perfect method'), a partial gesture for the missing consonant is often seen in the timing and formant movements on a spectrogram. It would be interesting to know the role of partial gestures in the production and perception of epenthesis. The gestures of epenthesis are more difficult to observe on spectrograms, however, since the consonants are often homorganic (thus eliminating formant movements) and nasals are notoriously difficult to distinguish. Ali et al. (1979) used recordings of oral pressure and oral and nasal airflow synchronized with the acoustic signal. Such a methodology shows promise for distinguishing the overlapping gestures of epenthesis.

## Acknowledgements

This article derives from research projects guided by Patricia Keating. Thanks also to Robert Port and an anonymous reviewer, whose suggestions led to substantial revision of the analysis.

## Appendix 1: Word list

'Say \_\_\_ again'.

1. consent
2. science
3. blunt saying
4. presents (noun)
5. census
6. intense
7. dents us
8. insurance
9. sensible
10. non-cents
11. conceal
12. incense (verb)
13. contents

14. sensory
15. offense
16. punts us
17. density
18. prescience
19. spent savings
20. defense
21. console
22. in tents
23. consist
24. presence

## Appendix 2: Design of the TIMIT corpus

There are 2342 sentences

- |   |  |                     |
|---|--|---------------------|
| A | 2 sentences for dialect identification   | (630 speakers each) |
| B | 450 sentences devised to place all American English phones in all possible segments contexts | (7 speakers each)   |
| C | 1890 sentences from the Brown corpus   | (1 speaker each)    |

There are 630 speakers. Each reads 10 of the sentences: 2 from set A, 5 from set B, and 3 from set C.

## References

- ALI, L., DANILOFF, R. & HAMMARBERG, R. (1979). Intrusive stops in nasal-fricative clusters: an aerodynamic and acoustic investigation. *Phonetica* **36**, 85–97.
- BARNITZ, J. G. (1974). Bloom-p-field, Chom-p-sky, and phonetic epen-t-thesis. *Studies in Linguistic Sciences* **4**, 1–13.
- BLANKENSHIP, B. (1992). What TIMIT can tell us about epenthesis. *UCLA Working Papers in Phonetics* **81**, 17–25.
- CLEMENTS, G. N. (1987). Phonological feature representation and the description of intrusive stops. In Bosch, A. et al. (eds.), *Parasession on autosegmental and metrical phonology*, 29–50. Chicago: Linguistic Society.
- DINNSEN, D. A. (1980). Phonological rules and phonetic explanation. *Journal of Linguistics* **16**, 171–191.
- FOURAKIS, M. & PORT, R. (1986). Stop epenthesis in English. *Journal of Phonetics* **14**, 197–221.
- LEE, S. (1990). The duration and perception of English epenthetic and underlying stops. Paper presented at the Fall 1990 Conference of the Acoustical Society of America.
- OHALA, J. (1974). Experimental historical phonology. In Andersen, J. M. & Jones, C. (eds.), *Historical linguistics II: Theory and description in phonology*, 353–389. Amsterdam: North-Holland Publishing Co.
- WARNER, N. & WEBER, A. (2001). Perception of epenthetic stops. *Journal of Phonetics* **29**, 53–87.
- ZWICKY, A. M. (1972). Note on a phonological hierarchy in English. In Stockwell, R. P. & Macauley, R. K. S. (eds.), *Linguistic change and generative theory*, 275–301. Bloomington: Indiana University Press.