The duration and perception of English epenthetic and underlying stops.

Sook-hyang Lee

Department of Linguistics, The Ohio State University

Abstract

In American English, an intrusive stop occurs before the fricative in words such as *tense* and *false*, making them very much like words with underlying stops, such as *tents* and *faults*. Ohala (1975) treats the inserted stop as an artifact of universal physiological or aerodynamic constraints. But this approach can't account for the fact that South African English speakers don't insert the stop between sonorant and fricative clusters (Fourakis and Port, 1986). Another approach posits a language- or dialect-specific phonological rule which inserts a phonological segment (Zwicky, 1972). Fourakis and Port (1986), argue against this approach on the ground that in some pairs the intrusive stop is significantly shorter than the underlying one (although the difference is always very small). This paper presents perception data and duration measurements supporting something like Zwicky's approach. Phrases with intrusive and underlying stops (*intense* and *in tents* , respectively) in citation forms produced by three speakers of Mid-Western dialects were presented over earphones in random order for subjects to identify. Identification was very poor, just at chance level. Also, duration measurements of the silence gap between the /n/ and /s/ in these words show no significant difference, contrary to Fourakis and Port's findings. Moreover, token judgments in the perception experiment show very poor correlation with the durations except for one speaker, implying that whatever duration differences there might not be a crucial cue that listeners exploit for labeling the words with epenthetic and underlying stops.

Introduction

The mapping between discrete phonological representations and the continuous speech event has been a controversial subject in phonetics. The occurrence of epenthetic stops in American English is one phenomenon in which there may be a discrepancy between the two representations. As shown in Fig. 1, in American English, an epenthetic stop occurs before the fricative in words such as dense and false, making them very much like words with underlying stops, such as dents and faults. Fourakis and Port (1986) measured the durations of these silent gaps in a cross-dialect study, and found, first, that in South African English, the epenthetic stops do not occur, and second, that in American English their durations are very slightly shorter than those of underlying stops.

Zwicky treats the epenthetic stop as the output of a dialect-specific phonological insertion rule. For example, in the case of *dense*, the silent gap between /n/ and /s/ is produced by inserting /t/ between them in the syllable coda position. On the other hand, Ohala treats the epenthetic stop as an artifact of universal physiological or aerodynamic constraints. According to him, the velum moves quickly compared to the tongue. Therefore, for *dense*, the velum raises for the following fricative /s/ before tongue tip closure for the nasal /n/ is released.

However, these two approaches have problems accounting for Fourakis & Port's findings: Zwicky's phonological approach cannot account for their finding that the duration of the epenthetic stop closure is significantly shorter than that of an underlying stop. Ohala's universal phonetic approach cannot account for their finding that in South African English, epenthetic stops do not occur.

In order to explain the significantly shorter duration of epenthetic stops, what they call "incomplete neutralization", Fourakis & Port suggest that the epenthetic stops are products of language specific phonetic rules, or phase rules, which have access to phonological structure. Ali et al. in an earlier study suggested three different ways in which the timing of gestures involved in /ns/ sequences could result in the appearance of an epenthetic stop.

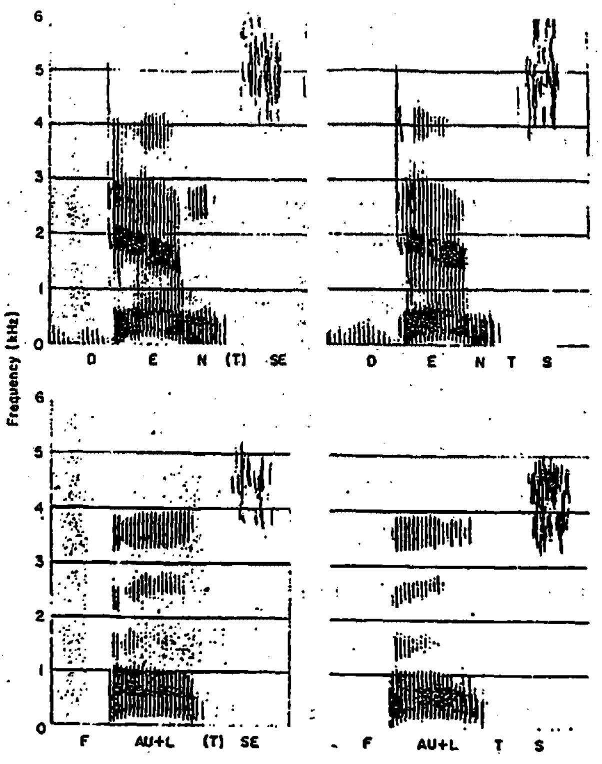

Figure 1. Spectrograms of American productions of *dense-dents* and *false-faults* from Fourakis and Port (1986).

If we think of the articulation of an /ns/ sequence as an arrangement of gestures as in Fig. 2, we can state Ali et al.'s three possibilities in terms of phase rules which affect one or another gestures. First, the tongue tip gesture might be phased late with respect to the other relevant gestures. This would result in a short oral stop articulation. Second, the velic aperture gesture might be phased late. The venting of the air through the nose might prevent the buildup of air pressure behind oral constriction necessary for the onset of turbulence. Third, the glottal gesture might be phased late. The lesser flow of air through the still adducted vocal folds might cause a similar delay in the buildup of air pressure necessary for [s] frication.

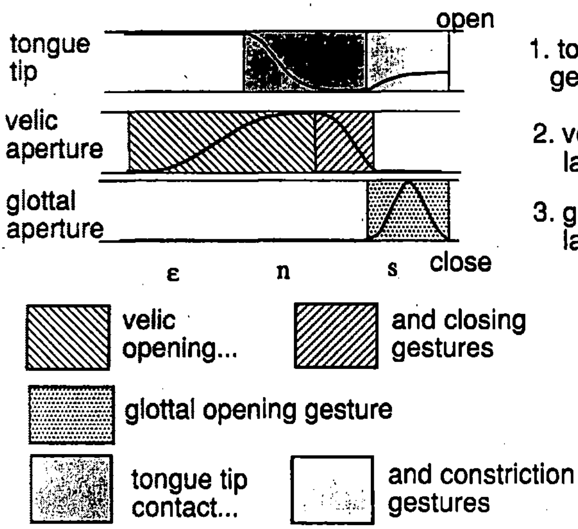

Figure 2. Ali et al.'s three possibilities of occurrence of epenthetic stops in American English stated in terms of phase rules.

An alternative explanation has been proposed by Clements (1987). He claims that the intrusive stops must be products of phonological rules. One of his arguments is that in dialects where underlying voiceless stops in the syllable coda trigger rules of glottalization, the intrusive stops do, too. Within the autosegmental phonology framework, he assigns intrusive stops different representation from underlying stops to account for "incomplete neutralization". According to Clements, the sequence /nts/ in *in tents* would have 3 C-slots while the sequence /ns/ in *intense* has 2 C-slots and the epenthetic stop is the result of spreading the oral cavity node of the nasal /n/ to the following fricative without delinking the node for the fricative as shown in Fig. 3 from Clements.

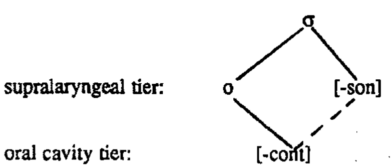

Figure 3. Oral cavity node spreading rule which derives epenthetic stops in /-ns/ sequences at syllable coda position in American English. [From Clements, 1987)

As a result, duration differences between underlying stops and epenthetic stops are automatically explained by their different surface forms. The underlying stop occupies its own C-slot as in Fig. 4a, whereas the epenthetic stop shares in a sense the first half of the /s'/ s C-slot as in Fig. 4b.

This paper reports on perceptual and durational evidence which might shed light on the phonetic and phonological status of epenthetic stops. If the underlying and epenthetic stops show a statistically significant difference and listeners can tell tokens of intense and in tents apart, the two stops are phonetically different. If there is no significant duration difference but listeners can distinguish them, then duration is not the salient perceptual cue and some other perceptual cue(s) might play a crucial role. If the duration difference between two types of stops is significant but listeners cannot tell the two tokens apart, then listeners are insensitive to the stop duration difference even though they are phonetically different. If there is no significant duration difference and listeners cannot tell them apart, we can say that they are phonetically the same. Finally, the correlation between stop duration and perception of epenthetic vs. underlying stops was calculated to see how much listeners depend on the stops' duration during the perception experiment.

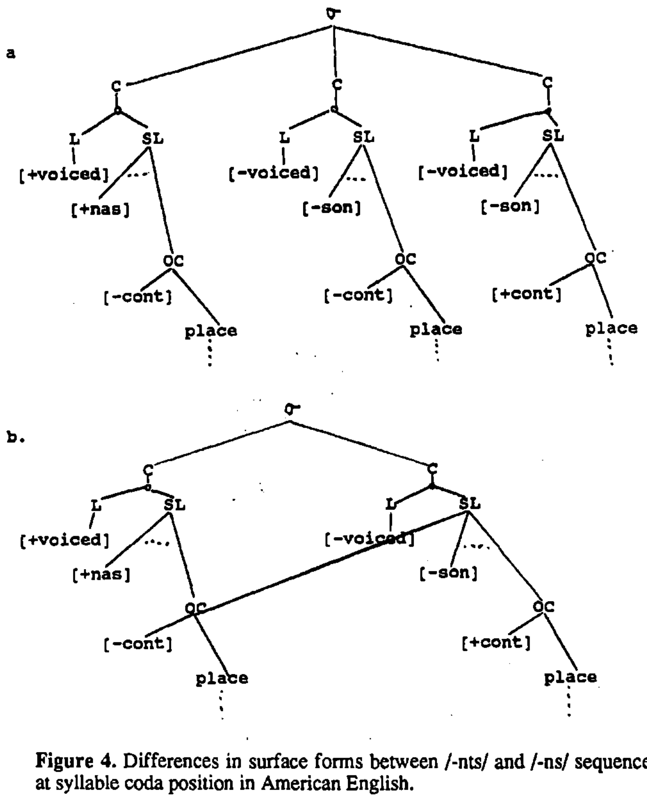

Figure 4. Differences in surface forms between /-nts/ and /-ns/ sequences at syllable coda position in American English.

Experiment 1: Acoustic duration of stops

Method

In the first experiment, duration measurements of segments, including epenthetic and underlying stops, were made. The corpus consisted of the 7 short dialogues shown in Table 1. In a sound-proof booth, one male and two female speakers of Mid-Western dialects read the corpus which had been prepared so that there were 6 repetitions of each type at 3 loudness levels in a randomized order. The subjects read each second sentence in a dialogue and then repeated the target word or phrase in citation form. For each type, this yielded 18 tokens of citation forms and 18 of noncitation forms per speaker.

Table I. Corpus

Noncitation forms	Citation forms
1. Isn't the light too glaring here? Yes, it's quite **intense** and bright.	**intense**
2. Aren't Mongolians nomads? Yes, they live and yurts.	**in tents**
3. Aren't the refuges in houses now? No, they're still in tent cities and shacks.	**in tent cities**
4. Isn't the population mostly in relaxed rural areas? No, they're still **in tense cities** and suburbs.	**in tense cities**
5. Hasn't the company reduced the number of branches? No, they're still **in ten cities** and four countries.	**in ten cities**
6. Isn't the company in sixteen cities now? No, they're only **in ten cities.**	**in ten cities**
7. Are these numbers the frequencies and everything? No, they're only **intensities.**	**intensities**

In this paper, I will concentrate on only intense and in tents. Duration measurements of all the segments were done using a spectrograph display. As shown in Fig. 5, the segmentation criteria were the same as for Fourakis and Port except that VOT was not included with the vowel but measured separately.

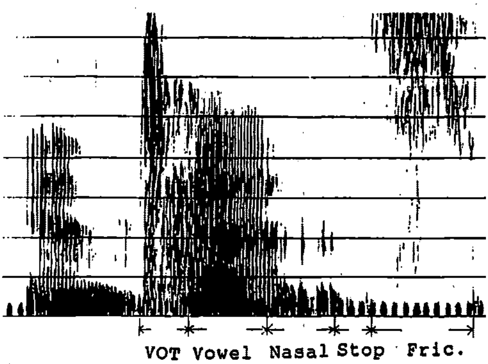

Figure 5. Segmentation of *intense* from spectrogram.

Results

As shown in Fig. 6, there is variation among speakers with respect to whether a silent gap appears: the two female speakers, MS and MC, always produced stops in both words but the other speaker ES did not. Sometimes there was no silent gap between /n/ and /s/ in *in tents* and the silent gap did not always appear in *intense*. This was true of citation forms as well.

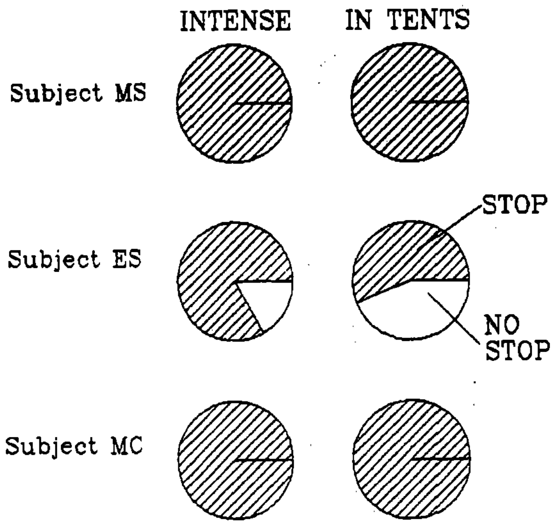

Figure 6. Percentage of occurrence of acoustic stops in the speech of each speaker in noncitation forms.

Looking at all of the segment durations for noncitation forms given in Fig. 7, speaker ES produced short epenthetic and short underlying stops compared to the other two speakers. Fig. 7 includes the token with no stop. However, even when these are excluded, speaker ES's stop duration is still much shorter.

An analysis of variance shows that in noncitation forms, the nasal, stop, and fricative of speaker MS in *in tents* were significantly longer than those in *intense*. The nasal of speaker ES in *in tents* was significantly longer than its counterpart in *intense*. (This result is counter to what would be expected if *tents* has underlying voiceless stop and *tense* does not, since the duration of vowel is negatively correlated with the number of consonants in the syllable coda.) The VOT of speaker ES and MC in in tents was significantly longer than that in *intense*. It may be because of the presence of a word boundary. That is, word initial /t/ in *in tents* may have longer aspiration as a signal of the word boundary.

Figure 8 compares citation and noncitation forms for subject MS. Citation and noncitation forms differed in two ways: first, in both words, all segments in citation forms are much longer than those in noncitation forms. (This was also true of the other speakers.) Second, in citation forms, none of the segmental duration in *intense* is significantly different from its counterpart in *in tents*. That is, although unlike the other speakers, speaker MS had a significant difference in stop duration in noncitation forms, this difference was reduced to the point of disappearing in citation forms, perhaps because of a different prosodic structure.

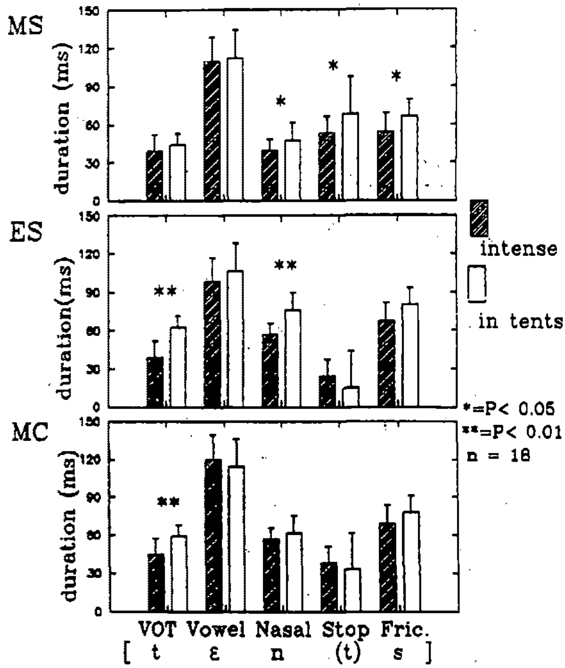

Figure 7. Duration measurements for each speaker in noncitation forms

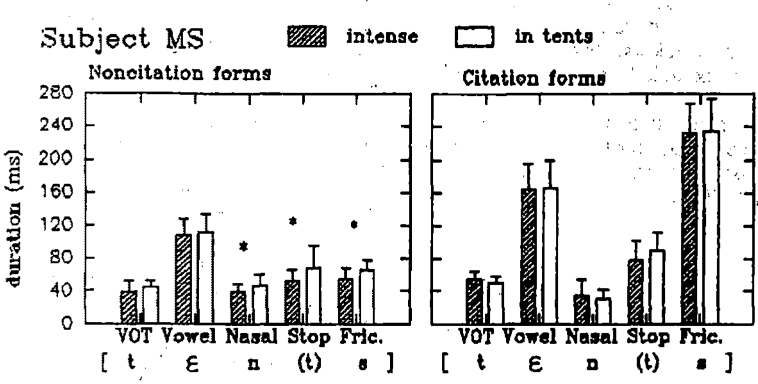

Figure 8.. Duration measurements for speaker MS.
*= P < 0.05, ** = P < 0.01, n = 18.

Discussion

The results of this experiment seem to suggest that *intense* and *in tents* might be the same phonetically for the speakers ES and MC, unless there is some factor other than duration which is used by listeners to distinguish them from one another. For speaker MS, on the other hand, there seems to be a small difference but only in noncitation forms.

Experiment 2: Perception Task

Method

In order to see whether listeners perceive a difference between *intense* and *in tents,* I ran a perceptual identification experiment. I only used citation forms in this experiment. I made a stimulus tape using the 18 tokens for each word type produced by each speaker for the first experiment. All the tokens for each speaker were repeated twice and two sets were constructed. In each set, all tokens were separated into two blocks and their order was randomized keeping the number of each word type constant within a block. Fifteen adult native speakers, three of whom were the subjects who produced the tokens, performed a forced answer identification test. For each speaker, the number of correct responses was averaged across all subjects.

Results

As shown in Fig. 9, listeners performed very poorly, essentially at chance level for all three speakers. Speakers even performed poorly for tokens which they produced themselves. Furthermore, even for the same token, listeners often responded differently to the first repetition and the second repetition when tokens were in different blocks.

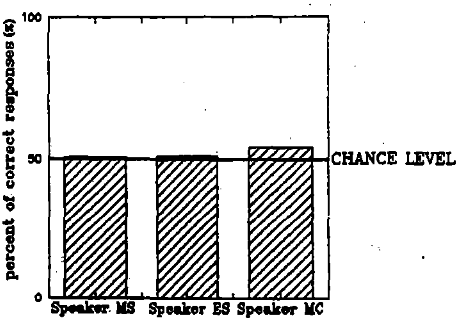

Figure 9. Percent of correct responses for identification task.

Discussion

The poor identification of tokens in this perception study demonstrates that these listeners could not tell *intense* and *in tents* apart and that they guessed when they were asked to label the tokens. These results and the acoustic duration measurement data together suggest that there are no crucial perceptual cues available for listeners to distinguish underlying and epenthetic stops from each other, and that the two types of stops in citation forms are not phonetically different. Even potential phonetic cues to the number of words constituting the tokens (one word for *intense* vs. two words for *in tents*) didn't help listeners distinguish the tokens from each other.

Experiment 3: Correlation of token judgments with stop durations

Method

Closer examination of the duration of stops in the citation tokens shows that the durations of epenthetic and underlying stops lie along two continua which overlap with one another, as shown in Fig. 10. Despite the overlap, speaker MS tended to have a longer stop duration in underlying stops, but this difference is not apparent in the other speakers.

It might be possible that listeners make use of the stop duration differences for speaker MS in the perception experiment even though they were not significant statistically. That is, generalizing from the small duration differences for this one speaker, listeners might tend to respond to shorter silent gaps as epenthetic and to longer gaps as underlying.

In order to test this possibility, I correlated perceptual responses with stop durations. If the stop duration difference is used as a perceptual cue, there should be a correlation between longer duration and more responses toward the underlying stop.

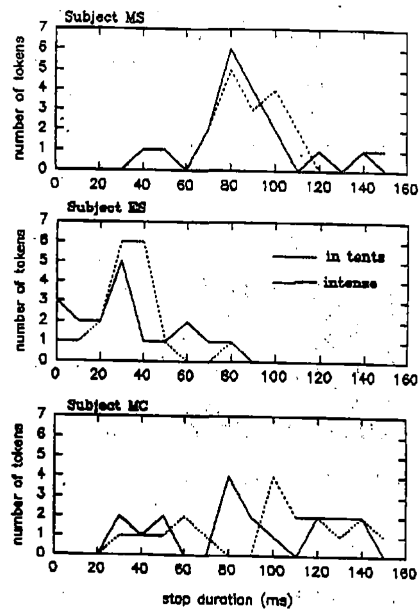

Figure 10. Range of duration of epenthetic and underlying /t/ produced by three speakers.

Results

As shown in Table 2, the correlation between the duration of stops in the tokens and subjects' responses to them is very poor except for one speaker, MS. The other two speakers, ES and MC, show even slightly negative correlations.

Table II. Correlation of token judgments in perception experiment with the duration of stops.

Speaker MS	0.55 ($P < 0.0005$)
Speaker ES	-0.095
Speaker MC	-0.13

Discussion

The poor correlation in the cases of speaker ES and speaker MC would be another indication that the two tokens are phonetically the same for these speakers. For speaker MS, on the other hand, even though the correlation was not very high, it still explains some of the perceptual variance, 25%. This result seems to suggest that there were some cases where duration differences were salient enough for listeners to use as a perceptual cue for speaker MS.

Conclusions

In this paper, I tried to determine the phonetic and phonological status of the so-called epenthetic stops in English. Acoustic duration measurements and data from the perception, and correlation studies together demonstrate that epenthetic and underlying stops are phonetically the same for some subjects. The duration of the stops were not different and listeners could not tell the two tokens apart. Fourakis and Port's study also showed that for some lexical items, the two types of stops were phonetically the same. They found the significant difference for only one minimal pair, *dense* and *dents* but not other pairs such as *tense* and *tents*. The subject who did show a difference in this study showed it more clearly in noncitation forms: therefore, the difference is also context sensitive. In future studies, I will try to figure out with more subjects and more lexical items, what about the context produces the difference .

* The original version of this paper was presented at the 121st meeting of the Acoustical Society of America, Baltimore, Maryland, 1991.

References

Ali, L. H. & R. Hammarberg. 1979. 'Intrusive stops in nasal-fricative clusters: An aerodynamic and acoustic investigation', *Phonetica* 3: 85-97.

Beckman, M. 1984. 'Spectral and perceptual evidence for CV coarticulation in devoiced /si/ and /syu/ in Japanese', *Phonetica* 41: 61-71.

Browman, C. P. & L. Goldstein. 1987. 'Tiers in articulatory phonology, with some implications for casual speech', *Status Report on Speech Research* 92. Haskins Lab.

Browman, C. P. & L. Goldstein. 1989. 'Articulatory gestures as phonological units', *Phonology* 6: 201-251.

Clements, G. N. 1985. 'The geometry of phonological features', *Phonology*

Yearbook 2.
Clements, G. N. 1987. 'Phonological feature representation and the description of intrusive stops', in A. Bosch et al. (eds.) *Autosegmental and Metrical Phonology*. Parasession of the Annual Meeting of the Chicago Linguistics Society, Chicago, 1987.
Davis, S. 1989. 'The location of the feature [continuant] in Feature Geometry', *Lingua* 78: 1-22.
Fourakis, M & R. Port. 1986. 'Stop Epenthesis in English', *Journal of Phonetics* 14: 197-221.
Fowler, C.A. 1980. 'Coarticulation and theories of extrinsic timing', *Journal of Phonetics* 8: 113-133.
Hayes, B. 1986. 'Assimilation as spreading in Toba Batak,' *Linguistic Inquiry*, 17: 467-499.
Kim, H. S. 1990. *The Manner Features in Phonological Representations*, Ph.D. dissertation, University of Indiana.
Lee, S-H. 1990. 'Evidence from glottal articulation for the analysis of intrusive stops in English,' *Journal of the Acoustical Society of America* 87:Suppl.1, p.S121.
Nolan, F. 1990. 'The Descriptive Role of Segments: Evidence from Assimilation' in D. R. Ladd and G. Docherty ed., *Laboratory Phonology* 2 (papers from the second conference on laboratory phonology). Cambridge University Press.
O'Dell, M. & R. Port. 1963. 'Discrimination of word-final voicing in German', *Journal of the Acoustical Society of America* 73: suppl.1, p.S31.
Ohala, J. J. 1975. 'Phonetic explanation for nasal sound patterns', in Ferguson, Hyman, Ohala ed., *Nasalfest: papers from a symposium on nasals and nasalization*, 289-316, Language Universals Project, Stanford.
Ohala, J. J. 1981. 'Speech timing as a tool in phonology', *Phonetica*, 38: 204-212.
Ohman, S. E. 1966. 'Coarticulation in VCV utterances: Spectrographic measurements', *Journal of the Acoustical Society of America*, 39: 151-168.
Zwicky, A. M. 1972. 'Note on a phonological hierarchy in English,' in R.P Stockwell & R. K. S. Macauley, ed., *Linguistic Change and Generative theory*, 275-301.