CONTEXTUAL INFLUENCES ON DEVOICING OF /Z/ IN AMERICAN ENGLISH

Caroline L. Smith
University of California, Los Angeles, USA

ABSTRACT

The devoicing of /z/ by speakers of American English was examined in a variety of sentence contexts using acoustic, airflow and EGG data. Although speakers differed in overall frequency of devoicing, they showed similar rank orderings for frequency of devoicing in different contexts. Both the immediate phonological context of /z/ and the prosodic strength of its position in the word influence the likelihood of devoicing.

BACKGROUND

Speakers of English often do not fully voice obstruents that are phonologically categorized as voiced. Voiced fricatives are often considered to require particularly precise conditions in the vocal tract: subglottal pressure must be higher than oral air pressure in order to produce voicing, but oral air pressure needs to be higher than atmospheric pressure to produce turbulence at the supralaryngeal constriction [1]. The term "devoicing" is used here to describe an absence of vocal fold vibration in the production of sounds normally categorized as voiced.

A number of previous studies have shown that devoicing is common in voiced fricatives in both British [2, 3, 4] and American English [5, 6, 7]. This experiment uses instrumental data to investigate two questions about the voicing of /z/ in connected speech. (1) In what environments is /z/ most likely to be devoiced? (2) How does devoiced /z/ differ from /s/?

Mechanisms of devoicing

Most previous studies have used acoustic data to identify the presence or absence of voicing. The temporal characteristics of devoicing in fricatives have been documented extensively, particularly by [3] and [7]. A greater likelihood for devoicing a voiced fricative when it is adjacent to a voiceless sound or silence has also been noted, suggesting that a kind of assimilation in voicing state is at work ([3, 4, 6, 7]). There is less information available on the physiological mechanisms involved in devoicing, such as whether devoicing is the consequence of a controlled opening movement of the glottis or is a passive consequence of the aerodynamic conditions that Ohala [1] suggests make voiced fricatives difficult to produce.

There is some evidence that the glottis does open during devoiced fricatives. Haggard [2] concludes, on the basis of F0 fall such as occurs following voicelessness, that the glottis opens during a voiced fricative in which glottal vibration ceases and then is re-initiated. In a study using transillumination [5], the majority of tokens of voiced fricatives showed evidence of glottal opening, whereas voiced stops mostly did not.

DATA COLLECTION

The present experiment was designed to investigate the devoicing of /z/ in a variety of phonological environments in natural speech. Speakers of American English produced 4 to 6 repetitions of 19 sentences. In these sentences, /s/ and /z/ occurred in contexts matched for type of neighboring sounds and position in word or phrase; the matched pairs of /s/ and /z/ occurred in different sentences.

Speakers wore a pneumotachographic mask to measure airflow and an electroglottograph (EGG) to measure vocal fold contact. These signals and the acoustic signal from a head-mounted microphone were recorded directly to disk at an 8000 Hz sampling rate. The airflow and EGG signals were low-pass filtered at 1000 Hz, the acoustic signal at 3000 Hz. A tape recording was also made, and digitized at 20000 Hz for acoustic analysis. Data from three speakers is reported here.

The EGG signal was used to identify where voicing was present in fricatives. The amplitude of one EGG cycle (maximum – minimum during one excursion) was measured at time of maximum RMS energy in the vowel preceding the fricative. The fricative was considered voiced during the portion of its duration that the amplitude of the EGG cycles exceeded one-tenth of the EGG cycle amplitude at the time of maximum energy in the preceding vowel. Voicing was considered to cease when the amplitude of an EGG cycle fell below this criterion. For each token of /z/, the percentage of fricative duration with voicing was calculated by dividing the duration of frication during which the EGG amplitude exceeded criterion by the total duration of acoustic frication.

The tokens of /z/ were categorized according to the percentage of their duration during which there was voicing. The three categories were:

0 - 25% voicing	devoiced
26 – 90% voicing	partially devoiced
90 - 100% voicing	voiced

Each category was analyzed separately and its acoustics and aerodynamics were compared with tokens of /s/ produced in matched phonological contexts.

DIFFERENCES BETWEEN /S/ AND /Z/

For each speaker, each category of tokens of /z/ was compared to an equivalent number of tokens of /s/ matched for phonological context. Paired t-tests show that /z/ and /s/ differ significantly in the following ways. For all speakers, the acoustic duration of frication is significantly shorter for /z/ than for matched /s/ for all groups of /z/. The acoustic duration of a vowel preceding /z/ is longer than a vowel preceding /s/ for all speakers when /z/ is partly or fully devoiced. However, for vowels preceding voiced /z/, only Speaker 1 had significantly longer durations; for Speakers 2 and 3 the durations of the vowels preceding voiced /z/ were not significantly different from vowels preceding matched tokens of /s/.

Measures of airflow also differed between /s/ and /z/. For Speakers 1 and 3, the mean airflow and the maximum airflow were lower for all groups of /z/'s than for the matched tokens of /s/. This was also true for Speaker 2 for partly or fully devoiced tokens of /z/; however, for this speaker there was no significant difference in the airflow measures between voiced /z/'s and matched /s/'s.

The differences in airflow between /s/'s and devoiced /z/'s could be due to a narrower glottal constriction in [z̥] than in [s], supporting the suggestion by Laver [4] that devoiced sounds may use a phonation type intermediate between the approximated vocal folds suitable for voicing and a fully open glottis. However, the substantial differences in duration between /s/ and all types of /z/ suggest that speakers are distinguishing the two sounds not only by phonation type and that, *contra* Laver, [s] and [z̥] should not be regarded as synonymous. Furthermore, given that mean airflow for devoiced [z̥] is comparable to that for [z], at least for Speakers 1 and 2, it seems unlikely that devoicing results from active widening of the glottis. Rather, it may be the consequence of a lower level of pulmonic effort.

FREQUENCY OF DEVOICING IN DIFFERENT CONTEXTS

Of the three speakers investigated so far, Speaker 1 was the least likely to devoice and Speaker 2 the most likely. Speaker 3 had the most tokens with partial devoicing. In the graphs below, the lightest shading corresponds to devoiced tokens of /z/. The darker gray corresponds to partially devoiced tokens, and the black to voiced tokens.

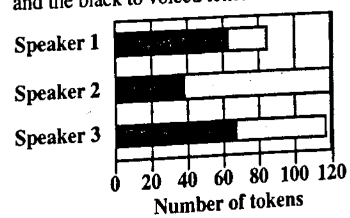

Figure 1. The number of tokens of /z/ in each of the three voicing categories: devoiced (□), partially devoiced (■), and voiced (■).

As for the influence of phonological context on the likelihood of devoicing, the known effect of the sound following a fricative [6] was confirmed in this experiment. The likelihood of devoicing by the different speakers for the /z/ at the end of "falls" is shown in Figure 2. The likelihood of devoicing differs considerably depending on whether a vowel, a voiced stop, a voiceless stop, or silence follows the /z/. (Only the relevant portion of each sentence appears in the graph labels.) Devoicing is least likely when the /z/ is followed by a vowel (the top bar in the graph) and most likely at the end of a sentence (the bottom bar). Although the speakers differ in how often they devoice overall, they all show a similar rank ordering among contexts.

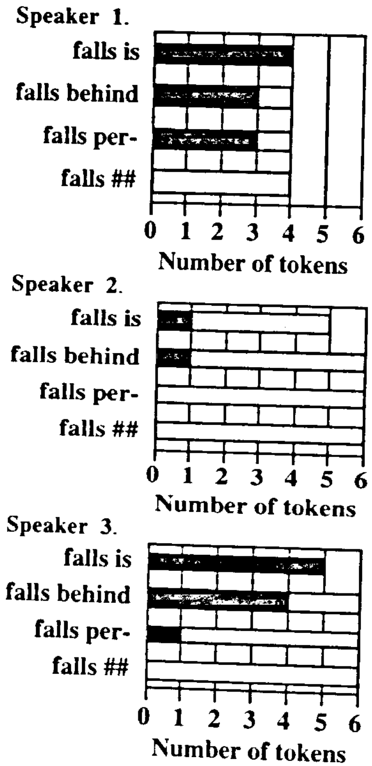

Figure 2. Number of tokens of /z/ at the end of "falls" that speakers produced as devoiced (□), partially devoiced (■), and voiced (■).

A similar pattern was observed for productions of word-final /z/ in "pause" with different following environments. This set of comparisons included the word "paused", in which the /z/ is in a syllable coda but is not word-final. In this coda position, 75% of the tokens of /z/ were devoiced and 25% partially devoiced, compared to 100% devoiced when "pause" was phrase-final. The /z/ in "pause" in all other contexts was less likely to be devoiced.

There is also some influence from the sound preceding a fricative. Syllable and word-initial /z/ were more likely to devoice when preceded by a voiced stop than a vowel. The top graph in Figure 3 shows more tokens with full voicing for syllable-initial /z/ preceded by a vowel ("de<u>ss</u>ert") than preceded by a voiced stop ("ob<u>s</u>erve"). The lower part of Figure 3 shows more tokens with full voicing for word-initial /z/ preceded by a vowel ("the <u>z</u>inc') than preceded by a voiced stop ("red <u>z</u>inc").

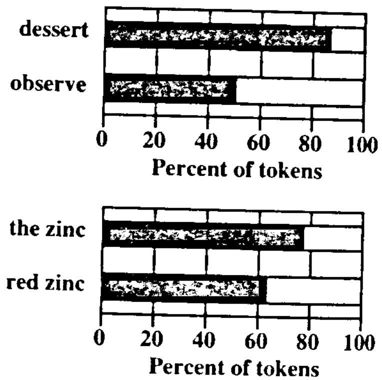

Figure 3. Effect of preceding vowel and voiced stop on likelihood of devoicing in syllable (top) and word (bottom) initial /z/.

All tokens of fully voiced /z/ in "observe" were produced by Speaker 1, the speaker who most often produced voiced /z/'s. All tokens of devoiced intervocalic /z/ in "dessert" and "the zinc" were produced by Speaker 2, the speaker most likely to devoice. Although devoicing was less frequent for the syllable- and word-initial /z/'s shown in Figure 3 than for the word-final /z/'s in Figure 2, nonetheless there were numerous tokens of initial /z/ with at least partial devoicing.

Stress also appears to play a role in determining the likelihood of devoicing a fricative, although the data are too sparse to make firm conclusions. Figure 4 illustrates the greater frequency of devoicing in word-final /z/ at the end of an unstressed syllable in "Dodgers" than at the end of a stressed syllable in "recurs". In both cases the target word was followed by a stressed syllable with initial voiced stop.

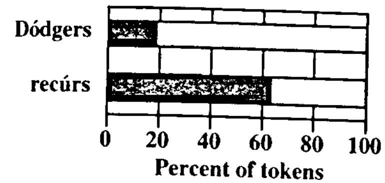

Figure 4. Percent of word-final /z/ that are devoiced at the end of an unstressed syllable (top) and a stressed syllable (bottom).

The comparisons presented here show that devoicing is more likely in positions that are generally the targets of lenition processes — in unstressed syllables, as part of a syllable coda, and at the end of a word or sentence [8]. It is not just the voicing characteristics of the immediate environment that condition the voicing of the fricative. The prosodic strength of the position in which the /z/ occurs is also very important in determining whether or not it will be voiced.

CONCLUSION

Different speakers vary as to how likely they are to devoice /z/. However, they shared similar rank ordering for frequency of devoicing in different phonological contexts. Speaker 2 rarely produced /z/ with glottal vibration during much of its duration, but nonetheless was more likely to produce at least some glottal vibration in those contexts that seemed to favor voicing.

Devoicing is most prevalent in precisely those environments where articulatory effort tends to be weaker. This pattern favors the interpretation that devoicing is a passive rather than an active process: speakers are not generating sufficient airflow from the lungs to maintain the trans-glottal pressure drop needed to maintain voicing. Much as Beckman et al. [9] model prosodic structure for temporal effects in production in terms of "sonority-time space", the occurrence of devoicing could be modeled in terms of the strength of a fricative's prosodic environment. In weaker prosodic environments, speakers may use lower airflow, resulting in a greater likelihood of devoicing. Faced with the voiced fricative 'dilemma' of maintaining a pressure drop across the supralaryngeal constriction to preserve the frication and a pressure drop across the glottis to preserve voicing, speakers of American English apparently prefer to maintain the frication.

ACKNOWLEDGMENT

This work was supported by grant NIH DC 00008 to the UCLA Division of Head & Neck Surgery.

REFERENCES

[1] Ohala, J. (1983), "The origin of sound patterns in vocal tract constraints", in P. MacNeilage (ed.), *The Production of Speech*, New York: Springer-Verlag, 189-216.

[2] Haggard, M. (1978), "The devoicing of voiced fricatives", *Journal of Phonetics 6*, 95-102.

[3] Docherty, G. (1992), *The Timing of Voicing in British English Obstruents*, Berlin: Foris.

[4] Laver, J. (1994), *Principles of Phonetics*, Cambridge: Cambridge University Press.

[5] Lisker, L., Abramson, A., Cooper, F. & Schvey, M. (1969), "Trans-illumination of the larynx in running speech", *JASA 45*, 1544-1546.

[6] Veatch, T. (1989), "Word-final devoicing of fricatives in English", presented at the Linguistic Society of America meeting, Washington, D.C.

[7] Stevens, K., Blumstein, S., Glicksman, L., Burton, M. & Kurowski, K. (1992), "Acoustic and perceptual characteristics of voicing in fricatives and fricative clusters", *JASA 91*, 2979-3000.

[8] Bell, A. & Hooper, J.B. (1978), "Issues and evidence in syllabic phonology", in A. Bell & J.B. Hooper (eds.), *Syllables and Segments*, Amsterdam: North-Holland, 3-24.

[9] Beckman, M., Edwards, J. & Fletcher, J. (1992), "Prosodic structure and tempo in a sonority model of articulatory dynamics", in G. Docherty & D. Ladd (eds.), *Papers in Laboratory Phonology II: Gesture, Segment, Prosody*, Cambridge: Cambridge University Press, 68-86.