UNCLASSIFIED
SECURITY CLASSIFICATION OF THIS PAGE

DTIC FILE COPY

REPORT DOCUMENTATION PAGE

AD-A199 923

1b. RESTRICTIVE MARKINGS

3. DISTRIBUTION/AVAILABILITY OF REPORT
Approved for public release; distribution unlimited.

4. PERFORMING ORGANIZATION REPORT NUMBER(S)

5. MONITORING ORGANIZATION REPORT NUMBER(S)
AFOSR-TR- 88-0936

6a. NAME OF PERFORMING ORGANIZATION
Harvard University

6b. OFFICE SYMBOL (If applicable)

7a. NAME OF MONITORING ORGANIZATION
AFOSR/NL

6c. ADDRESS (City, State, and ZIP Code)
Dept of Psychology
33 Kirkland St.
Cambridge, MA 02138

7b. ADDRESS (City, State, and ZIP Code)
Bldg. 410
Bolling AFB, DC 20332-6448

DTIC ELECTE OCT 05 1988 E

8a. NAME OF FUNDING/SPONSORING ORGANIZATION
AFOSR

8b. OFFICE SYMBOL (If applicable)
NL

9. PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER
AFOSR-87-0305

8c. ADDRESS (City, State, and ZIP Code)
Bldg 410
Bolling AFB, DC 20332-6448

10. SOURCE OF FUNDING NUMBERS

PROGRAM ELEMENT NO.	PROJECT NO.	TASK NO.	WORK UNIT ACCESSION NO.
61102F	2313	A4	

11. TITLE (Include Security Classification)
Context Effects in Recognizing Syllable-Final/z/ and /s/ in Different Phrasal Positions.

12. PERSONAL AUTHOR(S)
Peter C. Gordon

13a. TYPE OF REPORT	13b. TIME COVERED	14. DATE OF REPORT (Year, Month, Day)	15. PAGE COUNT
1st Annual Report	FROM 6/15/87 TO 6/15/88	Sept 6, 1988	36

16. SUPPLEMENTARY NOTATION

17. COSATI CODES			18. SUBJECT TERMS (Continue on reverse if necessary and identify by block number)
FIELD	GROUP	SUB-GROUP	speech perception, prosody, context effects phonetic segments, fricatives
05	09		

19. ABSTRACT (Continue on reverse if necessary and identify by block number)

Two experiments are reported that use gating methodology to examine the role of non-semantic aspects of sentential context in the recognition of phonetic segments. Performance in recognizing syllable-final /s/ and /z/ improves when the syllables are presented to listeners in sentential context as compared to when they are presented in isolation. It appears that listeners are able to use sentential information in order to factor out prosodically based variations in the temporal characteristics of speech in order to more accurately interpret durational cues to segment identity. These findings extend previous results on rate-dependent processing of overall speaking rate to the processing of local speaking rate, and they provide further demonstration of the importance of extended phonetic context in speech recognition.

20. DISTRIBUTION/AVAILABILITY OF ABSTRACT
☐ UNCLASSIFIED/UNLIMITED ☐ SAME AS RPT. ☐ DTIC USERS

21. ABSTRACT SECURITY CLASSIFICATION
UNCLASSIFIED

22a. NAME OF RESPONSIBLE INDIVIDUAL
Dr. Alfred R. Fregly

22b. TELEPHONE (Include Area Code)
(202) 767-5021

22c. OFFICE SYMBOL
NL

DD FORM 1473, 84 MAR

83 APR edition may be used until exhausted.
All other editions are obsolete.

SECURITY CLASSIFICATION OF THIS PAGE
UNCLASSIFIED

88 10 5 180

Statement of Work

The purpose of this work is to explore whether human speech recognition involves analyzing the speech signal at multiple levels of detail. In particular, it examines whether human listeners process relatively coarse-grained aspects of the signal in order to obtain information about the temporal structure of speech. Information about temporal structure, obtained in this way, could serve two major computational purposes. First, it could provide an independent basis for interpreting context-dependent durational cues to phonetic segments. Second, it could provide a computationally efficient way of allocating attentional resources to the temporal location of communicatively significant portions of the signal. The proposed research consists of two series of studies investigating these possible functions of coarse-grained information.

Processing of Phonetically Significant Durational Cues

The first study explores the extent to which heavily filtered speech conveys enough temporal information to induce rate-dependent processing of durational cues to phonetic categories. The second study generalizes the results of Study 1 to non-citation form speech, and to absolute identification tasks. Study 3 examines possible acoustic bases for coarse-grained cues to temporal structure, exploring the amplitude envelope of speech, various bands of low-frequency energy, and the role of broadband energy. Study 4 determines the amount of temporal detail preserved by various coarse-grained representations of speech. Study 5 examines factors influencing whether or not coarse-grained prosodic information is integrated with detailed phonetic information during speech recognition. Study 6 explores whether coarse-grained representations of speech convey information about local rate of articulation within a sentence.

Prosodic Influences on Attention

The first study in this series (Study 7) examines whether the allocation of attention to stressed syllables can be controlled by coarse-grained aspects of speech. Study 8 seeks to determine whether the expected advantage in processing stressed syllables, over unstressed syllables, results from their greater temporal predictability, as conveyed by coarse-grained representations of speech. Study 9 explores whether certain subtleties of the rhythmic structure of speech are extractable from coarse-grained aspects of speech. Study 10 focuses on whether a prosodically-induced characterization of a syllable as stressed, or unstressed, affects the strictness of the criteria used in segmental classification.

Status of Research

The body of the following document describes research that has been accomplished on the first section of the project as described in the Statement of Work (previous page). It is particularly relevant to the objectives of Studies 1 and 6 outlined in the statement of work.

Research Articles

Gordon, P.C. (1988). Induction of rate-dependent processing by coarse-grained aspects of speech. ***Perception & Psychophysics*, 43**, 137-146.

Gordon, P.C. (in press). Perceptual-motor processing in speech. In T.G. Reeve & R.W. Proctor (Eds.), ***Stimulus-Response Compatibility: An Integrated Perspective.*** North Holland.

Gordon, P.C. Context effects in recognizing syllable-final /z/ and /s/ in different phrasal positions. Manuscript to be submitted for publication, probably to the ***Journal of the Acoustical Society of America***.

Presentations

Induction of rate-dependent processing by coarse-grained aspects of speech. Haskins Laboratories, March 31, 1988.

Context Effects in Recognizing Syllable-Final /z/ and /s/ in Different Phrasal Positions

Peter C. Gordon

Harvard University

Running Head: Phonetic Context and Phrase Position

Address:

Peter C. Gordon
Department of Psychology
Harvard University
William James Hall
33 Kirkland St.
Cambridge, MA 02138

Abstract

Two experiments are reported that use gating methodology to examine the role of non-semantic aspects of sentential context in the recognition of phonetic segments. Performance in recognizing syllable-final /s/ and /z/ improves when the syllables are presented to listeners in sentential context as compared to when they are presented in isolation. It appears that listeners are able to use sentential information in order to factor out prosodically based variations in the temporal characteristics of speech in order to more accurately interpret durational cues to segment identity. These findings extend previous results on rate-dependent processing of overall speaking rate to the processing of local speaking rate, and they provide further demonstration of the importance of extended phonetic context in speech recognition.

The role of context in the perception of speech has been extensively studied and debated. Numerous experiments have documented the contribution of top-down semantic influences on lexical identification (e.g., Cole & Rudnicky, 1983; Samuel, 1981) and of adjacent segmental information on the phonetic interpretation of acoustic cues (e.g., Liberman, Coope. Shankweiler & Studdert-Kennedy, 1967). In addition, there has been an increase in understanding of the contribution of *extended phonetic context* to speech recognition, particularly with regard to speaking rate (Miller, 1981; in press) and syllable stress (Cutler & Norris, 1988; Grosjean & Gee 1987). The present paper seeks to extend that understanding by examining the role of non-semantic aspects of sentential context in recognizing the distinction between /s/ and /z/ in syllable-final position.

As implied above, "extended phonetic context", in the present use, should be understood in contrast to two other kinds of context: local phonetic context and semantic context. Local phonetic context consists of the coarticulatory information present in adjacent segments and coarticulatory effects on a segment from adjacent segments. "Semantic" context is present when high-level interpretation of an incoming message makes a specific lexical interpretation more plausible to a listener on semantic, pragmatic, or syntactic grounds. Extended phonetic context, in contrast, could convey a variety of information that is useful in segment recognition but that does not provide information about the specific articulatory gestures used to produce the segment or about the specific word that is being spoken. While a substantial amount of evidence has been gathered concerning the operation of local phonetic context and semantic context, evidence concerning the workings of extended phonetic context has been more elusive.

One area in which extended phonetic context might be taken to be important would be in normalizing acoustic cues to segments (particularly vowels) based on

speaker identity. Variability in some acoustic correlates of phonetic segments as a function of differences in speaker vocal tract have been well documented (Peterson & Barney, 1952), some effects of precursive sentential context on interpretation of vocalic information have been found (Ladefoged & Broadbent, 1957) and inter-speaker differences have been a major source of difficulty in developing machine models of speech recognition (Klatt, 1977). However, other experimental studies have not shown that human listeners have difficulty in recognizing speech sounds due to vocal tract differences, even when these sounds are gated out of their surrounding context (Verbrugge, Strange, Shankweiler & Edman, 1976). Instead, it appears that, at least for the recognition of vowels, dynamic cues to segment identity exist that make up for the variation in acoustic structure that exists due to speaker differences (Strange, Verbrugge, Shankweiler, & Edman, 1976).

Prosodic patterns provide a second, non-local source of variation in the acoustic structure of phonetic segments. Prosody influences the formant structure of vowels, fundamental frequency (F0), and the segment duration. These same acoustic characteristics all contribute, with varying degrees of strength, to the perception of a variety of phonetic segments. The prosodically relevant aspects of this variation must be factored out from the segmentally relevant sources of variation in order for accurate segment recognition to take place.

While there are some data available about listeners' use of prosodic context in interpreting acoustic cues to segment identity, not much is directly known about the effects of extended prosodic context on the interpretation of temporal cues to segment identity. However, variations in the temporal characteristics of segments due to prosodic patterns, such as phrase-final lengthening, can be considered as variations in local speaking rate (Klatt, 1976) and perhaps can be understood in relation to studies of overall speaking rate. In these studies (e.g., Gordon, 1988; Miller, Green & Schermer,

1985; Port, 1979; Summerfield, 1981) it has been found that the overall speaking rate of a precursive phrase can influence the interpretation of durational cues to segment identity. For example, Summerfield (1981) found that the boundary between voiced and voiceless percepts on a voice-onset time continuum became shorter as the speaking rate of a preceding phrase increased. Presumably this *rate-dependent processing* occurred because the temporal component of the voice-onset time was interpreted relative to the speaking rate of the preceding phrase (however, see Summerfield, 1981 for an alternative explanation). When the rate was fast, a shorter voice-onset time would indicate a voiceless segment.

The present research explores whether variations in local speaking rate produce effects on segment recognition similar to those found due to variations in overall speaking rate. Such an exploration is of interest because the patterns that convey information about local speaking rate are more complex than those that convey information about overall speaking rate. Overall speaking rate, by definition, applies to an entire utterance. A single speaking rate parameter could be extracted from extended phonetic context, and used to adjust the interpretation of durational cues to segment identity. In contrast, in order to use extended phonetic context to adjust for local rate variations, a listener must extract information from one stretch of speech to predict a different speaking rate for another stretch of speech. The experiments to be reported examine whether listeners rely on extended phonetic context for this purpose in accounting for the durational effects of phrase-final lengthening on acoustic cues to segment identity.

The present studies also explore the usefulness of gating methodology for studying rate-dependent processing in segment recognition. The gating procedure involves comparing the accuracy of recognizing a unit of speech presented in context relative to the accuracy of recognizing that unit removed from context (Pollack &

Pickett, 1963; Grosjean, 1980). The present experiments will see if extended phonetic context is important in recognizing a segment on a syllable which contains durational variation due to a prosodic pattern. The gating procedure offers the advantage that its dependent measure, percent correct, has a more natural interpretation than the more common measure of contextual importance of examining boundary shifts in phonetic judgments of an acoustic continuum. Furthermore, the gating procedure involves less drastic manipulation of the acoustic cues to the segment in question than does a procedure such as creating an acoustic continuum. This lessens concerns that the obtained results might reflect stylized, or unnatural speech stimuli. Despite these advantages, the gating procedure, because it does not directly manipulate acoustic cues to segment identity, is less definitive than the boundary shift method in providing information about the precise acoustic cues whose interpretation is being influenced by the context. Thus, the present use of the gating procedure should be seen as complementing studies that have used the boundary-shift method to study rate-dependent processing in segment recognition.

Finally, in addition to providing information about the effects of extended phonetic context on segment recognition, which presumably generalize beyond the specific linguistic contexts and phonetic segments used here, it is hoped that the present study will contribute specifically to understanding the recognition of syllable-final /s/ and /z/.

Cues to Voicing in Syllable-Final /s/ and /z/

Syllable-final fricatives are often devoiced regardless of their voicing status phonologically (Denes, 1955; Soli, 1982). Thus, the presence of voicing during the frication portion can not reliably be used to distinguish /s/ from /z/ in syllable-final position. Instead, a number of temporal and spectral characteristics have been found to correlate with the distinction and to serve as perceptual cues; although the

consistency with which these cues are present, and their relative perceptual importance is not yet completely clear.

As with other syllable-final consonants differing only in phonological voicing, the distinction between /z/ and /s/ is reflected in the duration of the preceding vowel. Numerous acoustic-phonetic studies (e.g., Denes, 1955; Hogan & Rozsypal, 1980; Lehiste & Peterson, 1960) have shown that vowel duration is consistently longer in syllables ending with /z/ than in syllables ending in /s/; for example, by an average difference of 92 msec in Lehiste and Peterson (1960). Perceptual studies have shown that decreasing the duration of the vowel can shift listeners' identification of final consonants from /z/ to /s/ (Denes, 1955; Derr & Massaro, 1980; Hogan & Rozsypal, 1980; Soli, 1982). Thus, there is substantial evidence that vowel duration, or some psychological dimension with which it is correlated in these studies, serves as an important perceptual cue to the recognition of voicing in a subsequent fricative.

Vowel-offset duration provides a second correlate of the voicing of syllable final fricatives. Examination of spectrograms shows that the amplitude contour of the offset of voiced energy is more gradual before /z/ than /s/, and that /z/ contains a longer period of simultaneously present voicing and frication (Soli, 1982; Zue, 1985). While this provides acoustic-phonetic evidence of a possible role for vowel-offset characteristics in distinguishing between /z/ and /s/, there does not appear to be clear perceptual evidence of the importance of this cue. Soli (1982) found no difference between listeners' identifications of syllables in which these cues were present and syllables where the cues had been neutralized.

The duration of frication is a third temporal correlate of voicing status in syllable-final /z/ and /s/ (Denes, 1955; Umeda, 1977), with the voiceless /s/ usually having a longer period of frication than the voiced /z/. Evidence supporting a perceptual role for frication duration comes from studies showing that increasing the

duration of frication can shift listener's identifications from /z/ to /s/ (Denes, 1955; Derr & Massaro, 1980). However, Soli (1982) has countered that frication duration is a weak cue to voicing, suggesting that the studies showing an effect of frication duration on identifications have used very stylized synthetic stimuli, and that his own studies using more natural stimuli show only small effects of frication duration.

In addition to the above temporal cues it has also been argued that there are spectral and structural cues to the recognition of voicing in syllable-final fricatives. Soli (1982) performed acoustic-phonetic and perceptual studies on the role of "vowel structure" in distinguishing between the syllables /juz/ (as in the verb "to use") and /jus/ (as in the noun "the use"). In particular, he examined the relationship between the duration of the initial second formant transition and the steady-state of the vowel. He found that the proportion of the steady state (or conversely the proportion of the transition region) was predictive of the voicing of the final consonant, such that higher proportions of steady state were associated with /juz/ than with /jus/. Within the sample he studied, this stimulus characteristic provided a sufficient basis for discriminating between the two kinds of speech sounds. Perceptual studies showed that manipulating the steady-state proportion had a strong influence on listeners' identification of the speech sounds.

Natural Speech Sample

The utterances that were to serve as stimuli were selected so as to incorporate extra-segmental factors that would influence important temporal correlates of voicing in syllable-final /z/ and /s/. This was done by manipulating the phrasal position of the test syllable, the vowel of the test syllable and the initial consonant of the test syllable. The manipulations of phrasal position and of vowel were modeled after ones used by Luce and Charles-Luce (1985).

The four sentence frames, which manipulate two factors (phrase position and voicing status of the segment immediately following the test syllable), are shown in Table 1, and the twelve test syllables are shown in Table 2. Sentences 1 and 3 place the test syllable in a phrase internal position, whereas sentences 2 and 4 place them in a phrase final position. Therefore, the latter two sentences are subject to phrase-final lengthening (Klatt, 1976; Martin, 1970) in which the final syllable of a phrase has an increased duration. The increase in duration occurs across the various temporal cues (e.g., vowel and fricative duration) that can serve to distinguish /z/ and /s/. Using the absolute values of these temporal correlates as cues to the voicing discrimination is thus problematic because variation due to segment voicing and phrase position are conflated in the same duration. Knowledge of phrasal position, derived from surrounding context, may therefore be important in using these cues to correctly recognize voicing. The second factor, voicing of the segment immediately following the test syllable, influences the likelihood that the frication portion of the final /z/ or /s/ will be devoiced. A subsequent voiceless segment, as in Sentences 1 and 2, tends to result in devoicing of the preceding fricative, whereas a voiced segment, as in Sentences 3 and 4, is less likely to produce devoicing. Because the present focus is on devoiced segments, only Sentences 1 and 2 were analyzed and used as stimuli.

The vowels of the test syllables were /I/, /i/, and /a/. They were chosen in order to provide diversity among the test syllables, and because they differ in their inherent durations. For example, in the measurements of Peterson and Lehiste (1960) the vowel /I/ has an average duration of 180 msec, /i/ has an average duration of 240 msec, and /a/ has an average duration of 260 msec. This durational variation, like the influence of phrasal position discussed above, is conflated with the vowel duration cue to voicing of the final fricative. However, while information from a relatively long stretch of speech is necessary to determine phrasal position, the span of speech containing the vowel should provide the information necessary to determine vowel identity and hence

the expected duration of the vowel. This should be the case except for the possibility that given that vowel duration is a "secondary" cue for vowel recognition (House & Fairbanks, 1953), the vowels may be difficult to recognize when the syllables are gated from context (Verbrugge, et al., 1976). Therefore, perceptual ambiguity about the relevance of vowel duration for recognizing the voicing of the final fricative, while possible, seems less likely than ambiguity due to the influence of phrasal position.

The initial consonants of the test syllables were /b/ and /w/. They were selected to provide diversity in the sample of test syllables. In addition, syllables beginning with the semivowel /w/ are similar to the syllables beginning with /j/ studied by Soli (1982). The slow initial formant transitions in the /w/ provide a richer vowel structure than the faster transitions of the /b/ which may assist in recognizing the voicing of the final consonant. Furthermore, the duration of the formant transitions for /w/ are more likely to provide an adequate basis for normalizing rate variations than those of /b/, since Miller and Baer (1983) found that formant-transition durations for /w/ changed with speaking rate much more than those /b/.

The four sentence frames, two initial consonants and three vowels, when combined with the two syllable final consonants (/z/ and /s/) produced 48 utterances. They were spoken by three female laboratory assistants who were naive to the purposes of the experiment. Speaker CN grew up in northern New Jersey, speaker SL in suburban Maryland, and speaker SK in West Virginia. Each of the speakers produced two repetitions of the 48 utterances in different random orders. Different randomizations were used for each speaker. The utterance to be produced were presented on a video monitor, and the speaker was asked to read them in a natural speaking voice. The readings were done in a sound attenuating chamber. The utterances were picked up by a Shure Model SM59 microphone and recorded by a Nakamichi BX-100 cassette deck.

Measurements and Segmentation. The recorded utterances were low-pass filtered at 9.4 KHz and digitized at a sampling rate of 20 KHz. The following locations in the utterance were then marked for purposes of measurement and segmentation: (1) The beginning of the utterance which was defined as the onset of visible or audible energy. (2) The onset of the test syllable. In the case of /b/, this was defined as the release of the stop closure; in the case of /w/, it was determined by a combination of the end of audible and visible frication from the preceding /z/ and the beginning of a rise in the amplitude envelope. (3) The onset of fricative energy which was determined by the onset of high frequency, aperiodic energy. (4) The offset of voicing, which was determined by the disappearance of low frequency, periodic energy. (5) The end of the test syllable which was defined as the onset of the closure for the following stop consonant. And (6) the end of the utterance as indicated by the end of acoustic energy. In addition, when the test syllable began with /w/ the end of the second formant transition was measured. These locations were determined by examination of waveforms, listening to portions of waveforms, and examination of spectrographic displays.

Durational Analyses. The principal results of the analyses are shown in Table 3 which gives means and standard deviations of various durational measurements broken down by test syllable and phrase position.

Results for vowel duration are shown in the top section of Table 3. Actually, "vowel duration" is an accurate label only when /w/ is the initial consonant. In this case, the vowel duration is measured from the offset of the second formant transition to the onset of frication. In the case of /b/ as an initial consonant, the measurement is made from the release of the closure for the /b/ to the onset of frication. Thus "vowel duration" for syllables beginning with /b/ would be more properly referred to as consonant-vowel duration. However, this quantity will be used as a proxy for vowel

duration because of ease of measurement and because the duration of the formant transitions associated with initial /b/ have not been found to vary much as a function of speaking rate (Miller & Baer, 1983).

The expected effects on vowel duration can be seen in the relations between the various means. As a function of vowel identity, vowel duration increases from /I/ to /ɪ/ to /a/. Also, vowel duration is longer in phrase-final position than in phrase-internal position. Most importantly vowel duration is longer when the final consonant is /z/ than when it is /s/. This pattern holds between all /z/-/s/ pairs of mean duration when the contextual factors (consonant identity, vowel identity and phrasal position) are held constant. The pattern does not hold when the contextual factors are not taken into account.

Mean vowel-offset duration, shown in the second section of Table 3, varies very little as a function of the contextual variables. However, it differs considerably as a function of voicing of the syllable-final fricative, averaging 36.0 msec for /z/ and 14.4 msec for /s/. The vowel-offset means appear to offer a less context-dependent basis for discriminating /z/ and /s/.

Mean fricative duration, shown in section 3 of Table 3, also appears to provide some information in discriminating between /z/ and /s/. Fricative duration is longer for /s/ than /z/, although this difference is bigger in phrase-final position, 158 msec versus 96 msec, than it is in phrase-internal position, 117 msec versus 92 msec. (Alternatively, it could be stated that the effect of phrase-final lengthening is much greater for /s/ than for /z/.) Mean fricative duration did not vary much as a function of vowel identity (mean duration in msec: /I/ = 112; /a/ = 116; /ɪ/ = 117).

In addition to the above absolute durational measures, the relational measure of fricative/vowel ratio was also computed. This was done because it has been suggested

that fricative/vowel ratio may be the effective perceptual cue to the syllable-final /z/ - /s/ distinction (Denes, 1955), and because it has suggested that unitless, relational measures may in general offer a solution to the problem of rate variation in segment recognition (Port & Dalby; 1982; Soli, 1982). This ratio seems to discriminate fairly well between /z/ and /s/, if the specific context of the ratio is taken into account.

A second relational cue, which assesses vowel structure, is shown in section 5 of Table 3. The proportion transition is the proportion of the entire voiced portion of the syllable that is made up of the formant transitions into the vowel. This measure was proposed by Soli (1982) as a context-independent cue to distinguishing /juz/ and /jus/. The present analysis involves a different initial consonant, /w/ rather than /j/, and three different vowels. The measure does seem to distinguish /z/ from /s/ for the vowels /I/ and /i/, when vowel identity and phrasal position are taken into account. Its failure to distinguish /z/ and /s/ when the vowel is /a/ should perhaps be discounted because the low second formant of /a/ made it difficult to reliably measure the end of its initial transition. However, even in the case of the other vowels, this relational cue does not seem to provide a context-independent cue to voicing in syllable-final /z/ and /s/.

Quantitative Assessment of Diagnosticity of Durational Cues. The above review of the descriptive statistics of the various durational cues gives a qualitative sense of how the various measures could be used to distinguish between syllable-final /z/ and /s/. A quantitative measure of this ability was obtained by performing a series of discriminant analyses. These analyses indicate how well the optimal linear combination of these cues could do in classifying the tokens as containing the segment intended by the speaker.

The first analysis looked at vowel duration, vowel-offset duration, and fricative duration. All three variables significantly contributed to predicting fricative identity

(Chi Square (3) = 106.8, $p < .0001$). The standardized coefficients of the discriminant function are .796, .357, and -.549 for vowel-offset duration, vowel duration, and fricative duration respectively. The absolute values of these coefficients indicate the relative contribution of each variable to the prediction. Together, these variables successfully classify 89.6 percent of the tokens. A straightforward indication of the predictive value of each variable is given by the classification rate of each variable alone. Vowel-offset duration can be used to successfully classify 81.9 percent of the tokens (Chi Square (1) = 72.2, $p < .0001$). Classification based on vowel duration is successful in 61.8 percent of the cases (Chi Square (1) = 22.0, $p < .0001$). However, this figure is misleading since vowel duration is collapsed across initial consonant identity and the measures used in the two cases are not identical. When separate discriminant analyses are performed on syllables with different initial consonants some improvement in classification is obtained; 69.4 percent for initial /w/ (Chi Square (1) = 15.0, $p < .0005$) and 66.7 percent for initial /b/ (Chi Square (1) = 14.7, $p < .0005$). Classification success based on fricative duration was 62.5 percent (Chi Square (1) = 30.7, $p < .0001$).

Discriminant analyses also showed statistically significant classification success based on the relational variables of fricative-vowel ratio and proportion transition. For fricative-vowel ratio, the classification success rate is 73.6 percent for initial /b/ (Chi Square (1) = 27.7, $p < .0001$) and 79.2 percent for initial /w/ (Chi Square (1) = 27.2, $p < .0001$). For proportion transition, the classification success rate is 65.3 percent (Chi Square (1) = 6.4, $p < .02$).

In addition to the overall analyses, separate discriminant analyses using vowel duration, vowel-offset duration and fricative duration were conducted on the utterances produced by each speaker. These analyses yielded a success rate of 97.9 percent for the utterances of speaker CN (Chi Square (3) = 64.3, $p < .0001$), 95.8 percent for the utterances of speaker SL (Chi Square (3) = 42.2, $p < .0001$), and 81.3 percent for

speaker SK (Chi Square (3) = 16.3, $p < .005$). Further discriminant analyses on the utterances of speaker SK using the relational measures, and various combinations of the relational and absolute measures failed to disclose any basis for achieving a very high classification rate for the utterances of speaker SK.

The results of the discriminant analyses on single variables, both absolute and relational, indicate that none of these measures alone provides a context-independent basis for successfully classifying syllable-final /z/ and /s/. This finding limits the generality of a number of studies that suggest that these factors might provide a basis for recognizing /s/ and /z/ (Denes, 1955; Soli, 1982). Of these cues taken singly, vowel-offset duration and fricative-vowel ratio seem to other the best basis for classification.

The results of the discriminant analyses using the three absolute durational measures in combination show that a fairly high rate of successful classification can be achieved, especially for speakers SL and CN. Classification using these measures is essentially context-independent, since the classification process is given no information about the sentential position or the vowel identity. However, it is unclear whether listeners can encode the values of these durations as precisely as in the present analyses, if they encode them directly at all. Additionally, it is doubtful that listeners' decision criteria would be optimized as precisely for the present sample as they are in the discriminant analysis. On the other hand, it is of course possible that listeners might exploit additional cues, or configurations of cues, in recognizing syllable-final /s/ and /z/.

Experiment 1

The purpose of this experiment was to assess whether listeners' recognition of syllable-final /s/ and /z/ improves when the syllables are presented in sentential

context. Such a finding would indicate that listeners' interpretation of the voicing cues present in a syllable will, if possible, incorporate information from the extended phonetic context beyond the syllable, and that this incorporation leads to more accurate performance. In addition, the manner in which recognition of the fricatives depends on the various contextual factors, and the extent to which these interactions depend on the presence of context may shed light on the what cues are most important in recognizing syllable-final /s/ and /z/, and the role that context plays in interpreting those cues.

Method

Stimuli, Design and Procedure. The utterances described above served as stimuli in the experiment. Syllables were presented without context (gated) and in context. The gating was performed at the points described in the above section on measurements. All cuts in the waveform were made at zero crossings where the waveform was ascending.

The gating factor (gated vs. in context) was manipulated between blocks. Each block had 72 trials which contained one repetition of each of the factors (2 sentence positions x 2 initial consonants x 3 vowels x 2 final fricatives x 3 speakers). A block thus contained half of the utterances. Each subject listened to four blocks of stimuli, with the gating factor alternating between blocks. Half the subjects began with a block that contained gated syllables, and half began with a block that contained the syllables in context. The first two blocks that a subject heard contained the complete set of 144 utterances. Across subjects, the half of the utterances presented in the first block was counterbalanced.

Each block was divided into two sub-blocks of 36 trials, with a ten second rest interval between sub-blocks. The stimuli were output from the computer at 20KHz, low-pass filtered at 9.4 KHz, and recorded on cassette. The recordings were amplified

by a Rane HC 6 headphone amplifier and presented to the subjects over Sennheiser HMD 240 headphones at a comfortable listening level. Subjects were seated in a sound-attenuating chamber, instructed about the general procedure and asked to write "S" or "Z" on a response sheet depending on the final consonant of the test syllable.

Subjects. Twelve students, attending classes at Harvard University, served as volunteer subjects. They were all native speakers of American English, and none reported any known hearing disability. They were paid $4.00 for their participation in a single session that lasted approximately one-half hour.

Results

Table 4 summarizes the main findings. It shows the percentage of correct responses for /s/ and /z/ in the two phrasal positions. These percentages are shown on average as well as broken down according to the identity of the initial consonant of the syllable and of its vowel. Subjects made more correct identifications when the test syllables were presented in context than when they were gated, $F(1,11) = 10.9$, $p < .01$. There were no significant main effects of phrasal position [$F(1,11) < 1$], vowel identity [$F(2,22) = 3.1$, $p < .10$)], initial consonant identity [$F(1,11) = 1.6$, $p > .10$)] or of fricative identity [$F(1,11) = 3.1$, $p > .10$)].

There was a significant interaction between fricative identity and phrasal position; $F(1,11) = 107.5$, $p < .0001$. The form of this interaction was that more accurate identification of /s/ occurred in phrase-internal position than in phrase-final position, and more accurate identification of /z/ occurred in phrase-final position than in phrase-internal position. Furthermore, there was a significant three-way interaction between fricative identity, phrasal position, and whether or not the syllable was presented in context; $F(1,11) = 14.2$, $p < .005$. The form of this interaction was such that the beneficial effects of the presence of context were greater for /s/ in phrase-final

position than in phrase-internal position, and for /z/ in phrase-internal position than in phrase-final position.

In addition, there was a significant main effect of speaker; $F(2,22) = 98.4$, $p < .0001$. The mean correct identification were 93.0, 90.9 and 75.4 for utterances produced by speakers CN, SL and SK. This ranking is the same as that found in the discriminant analyses reported above. This suggests that the difficulty listeners had in identifying the utterances produced by speaker SK may be based in those stimulus characteristics that were employed in the discriminant analyses. It is not clear whether the tokens produced by speaker SK reflect her dialect, her idiolect, or some aspect of the manner in which she approached the recording task. More to the present concern, there was no significant interaction between speaker identity and presence of context, $F(2,22) < 1$. A number of other interactions involving speaker did achieve significance, however, the pattern behind them was unclear. An understanding of these interactions would require an exploration of the details of the speaker's idiolects that is beyond the scope of the present study.

Discussion

The results of the experiment clearly indicate that extra-syllabic contextual information is useful to listeners in recognizing syllable-final /z/ and /s/. Recognition accuracy improved by an average of 3.9 percent when syllables were presented in context, as compared to when they were presented in isolation. This suggests that the various cues to voicing contained within the syllables were not sufficiently context-independent so as not to benefit from the contextual information provided in the sentence frame.

The voiced fricative /z/ was identified more accurately in phrase-final position than in phrase-internal position, while the opposite relation was obtained for the

voiceless /s/. This finding suggests that the more important durational cues to voicing are present in the vocalic, rather than the fricative, portion of the syllable. The vocalic and fricative portions of syllables in phrase-final position are lengthened relative to those in phrase-internal position. Longer vocalic segments are associated with voiced percepts (/z/) while longer fricative segments are associated with voiceless percepts (/s/). Listeners' identifications thus appear to be more consistent with the information present in the vocalic than the fricative portion of the syllables. This finding is consistent with Soli's (1982) conclusion that fricative duration is a relatively weak cue for identifying syllable-final /z/ and /s/.

The extent of the influence of the presence of context depended on the interaction of fricative identity with phrasal position; the greatest contextual benefit occurred for /z/ in phrase-internal position (7.1 percent improvement) and /s/ in phrase-final position (6.3 percent improvement). This suggests that the presence of the context led to more accurate interpretation of the vowel-duration information. The most errors occurred for /z/ in phrase-internal position and /s/ in phrase-final position. It is likely that this is because the effects of phrase position on vowel duration were opposite to the effects of final-fricative voicing on vowel duration. The beneficial effects of the sentential context likely derived from its providing information about the phrasal position of the test syllable and thus a basis for factoring out some portion of variation in vowel duration that was extrasyllabic in origin.

Experiment 2

This experiment examines whether further understanding of the recognition of syllable final /s/ and /z/ varying phrase phrasal position can be achieved by examining listeners' identifications of the test syllables, both with and without context, after one of the major within-syllable cues to final fricative voicing has been eliminated. The discriminant analyses discussed above showed that vowel-offset duration provided

substantial context-independent predicative value about voicing in syllable-final /s/ and /z/. Its predicative value was greater than that of any other single cue measured in the syllables. Therefore, this stimulus characteristic should be considered as a candidate perceptual cue for recognizing voicing.

If vowel offset plays an important role in recognizing syllable-final /s/ and /z/, then its presence in the syllables used in the last experiment may have reduced the importance of some of the context-dependent cues present in the syllables. This, in part, would reflect the relative importance of the two kinds of cues. However, Wardrip-Fruin (1985) has shown that listening conditions can influence the importance of context-dependent cues relative to more context-independent cues. She found that the importance of vowel duration in cuing the voicing of syllable-final stop consonants increased relative to voicing during closure when the syllables were presented against a noise background. This is presumably because the low-amplitude voicing information present in the stop closure was less likely to be encoded under difficult listening conditions, allowing the more robust vowel duration information to carry the perceptual judgment. The vowel offset portion presently under study is similar to voicing during a stop closure in that its critical feature, the presence of voicing during frication, is low-frequency and low-amplitude. Therefore, it seems likely that this cue might not always be available to listeners as a basis for recognizing syllable-final /s/ and /z/. The present experiment attempts to simulate those conditions by editing out the vowel-offset portions of the syllables. If the vowel-offset characteristics are an important context-independent cue for recognizing syllable-final /s/ and /z/, then we may expect that the various contextual factors that vary in the stimulus set might have a greater influence in its absence.

While the above arguments suggest that vowel-offset may be an important perceptual cue, it must be recalled that Soli (1982) found little effect of vowel-offset

characteristics in his study of the /jus/ versus /juz/ distinction. Therefore, the present study offers further examination of the importance of vowel-offset duration *per se*.

Method

Stimuli, Design, and Procedure. The stimuli were the same as those in Experiment 1 except that the vowel-offset portion of the syllables was removed at the points described in the measurement section above. All the cuts in the waveform were made at ascending zero-crossings. The resulting syllables were free from clicks and sounded quite natural. Other aspects of the design and procedure were the same as in Experiment 1.

Subjects. Twelve new individuals, from the same pool as the previous experiment, served as paid subjects.

Results

The principal results are presented in Table 5 which shows the mean correct responses for /z/s and /s/s at the two phrasal positions presented in isolation and in context. These means are shown overall and broken down by the identity of the initial consonant and vowel of the syllables. As in the previous experiment, recognition accuracy was significantly better when the syllables were presented in context than in isolation; $F(1,11) = 15.4$, $p < .005$. Also, there were no significant main effects of vowel identity [$F(2,22) = 3.17$, $p > .05$], phrasal position [$F(1,11) = 3.24$, $p > .05$], or fricative identity [$F(1,11) < 1$]. In contrast to the previous experiment, there was a significant effect of initial consonant identity. Performance was more accurate for syllables beginning with /w/ than /b/ [$F(1,11) = 68.8$, $p < .0001$].

A significant interaction was obtained between phrasal position and fricative identity [$F(1,11) = 136.9$, $p < .0001$]. As in Experiment 1, /z/ was more accurately

identified in phrase-final position than in phrase-internal position, while /s/ was more accurately identified in phrase-internal position than in phrase-final position. Also, there was again a significant three-way interaction between phrasal position, fricative identity, and the presence of context [$F(1,11) = 10.5$, $p < .01$]. As in the previous experiment, improvement in recognition accuracy due to the presence of context occurred to the greatest degree for /z/ in phrase-internal position and /s/ in phrase-final position.

Some additional interactions, not found in the previous experiment, also emerged. The addition of context produced a greater improvement in accuracy in phrase-internal position than in phrase-final position [$F(1,11) = 5.6$, $p < .05$]. This interaction is probably best understood as a product of the beneficial effects of sentential context being mostly localized in phrase-internal /z/ (see the three-way interaction described above). The interaction of vowel identity and fricative identity was significant [$F(2,22) = 16.9$, $p < .0001$]: following the vowel /I/, /s/ was recognized 9.7 percent more accurately than /z/; following /i/, /s/ was recognized 6.7 percent more accurately than /z; and following /a/, /s/ was recognized 7.9 percent less accurately than /z/. This interaction is consistent with the idea that longer vowel durations are associated syllable-final /z/ as opposed to /s/. Vowel duration increases from /I/ to /i/ to /a/ and so does the percentage of /z/ responses. A significant interaction of vowel identity and the presence of context [$F(2,22) = 4.0$, $p < .05$] indicates that a greater context effect was found for /I/ than for /i/ or /a/.

As in Experiment 1, there was a significant main effect of speaker [$F(2,22) = 67.1$, $p < .0001$] with the utterances produced speaker CN were again recognized with the highest accuracy while those of speaker SK were recognized with the lowest accuracy. While there were several significant interactions involving speaker identity

and other factors manipulated in the experiment, there was no significant interaction of speaker identity and the presence of context [$F(2,22) < 1$].

Discussion

Overall recognition accuracy was 83.8 percent in Experiment 2 as compared to 86.4 percent in Experiment 1. However, recognition accuracy on /z/ decreased from 89.4 percent in Experiment 1 to 82.4 percent in Experiment 2 while recognition accuracy on /s/ improved from 83.4 percent to 85.2 percent. This suggests that the presence of simultaneous voicing and frication (the vowel-offset) is a moderately strong, though not overwhelming cue that a syllable-final fricative is voiced.

Removing the vowel-offset portion of the syllables caused additional effects of the phonetic environment of the fricative to emerge. Syllables beginning with /w/ were recognized more accurately than those beginning with /b/. The findings of Soli (1982) suggest that the higher recognition accuracy for syllables beginning with /w/ may be due to the vowel structure information present in the glide-vowel combination of the initial portion of the syllable. Furthermore, several interactions involving vowel identity were found. These interactions seemed to be best understood in terms of the relationship between vowel duration and syllable-final voicing.

With regard to the effects of the presence of sentential context, the results of Experiment 2 indicate that the improvement in recognition accuracy is quite specific. This improvement was 13.5 percent for /z/ in phrase-internal position. In this case the contextual information provided by the sentence frame is helping to overcome two factors that make accurate recognition difficult: the removal of the vowel-offset portion (which is a positive cue to a voiced status) and the short vowel duration due to phrase-internal position.

Conclusions

The results of the two experiments support the idea that extended phonetic context is of use to listeners in recognizing phonetic segments. The relatively large sample of speech that was measured and that provided the stimuli suggests that the findings are not due to idiosyncrasies of a particular model of the acoustic-phonetic makeup of syllables containing final /s/s or /z/s. The use of gating methodology indicates that the presence of context does not simply bias identifications but leads to actual improvement in recognizing the intended utterance of the speaker. While the gating methodology does not definitively specify the cues whose interpretation is influenced by context, the pattern of results provides a strong circumstantial case that the context allows more accurate interpretation of syllable-final voicing cues conveyed by vowel duration. This suggests that the sentential context provides information about the prosodically-based phrase-final-lengthening effect, and that listeners use this information in interpreting durational cues to segment identity.

While the experiments clearly indicate that the extended phonetic context, consisting of a sentence frame, is useful in segment recognition, they do not provide information about what aspects of the sentence frame convey the contextual information. Because the segments to be recognized were present on nonsense syllables, the contextual effect could not stem from semantic constraints on lexical identity. However, it is possible that recognition of the phrasal position of the test syllable was achieved through the syntactic constraints obtained by recognizing the surrounding words. If such a process is responsible, it would suggest that recognizing phonetic segments involves the interaction of very disparate levels of linguistic analysis. Alternatively, it is possible that the contextual effect derived from direct recognition of the phrasal position conveyed by the acoustic pattern of the sentence frame. Recognition of the phrasal position might make use of such acoustic patterns as the variations in the speech amplitude envelope and the F0 contour. A process such as this would be consistent with Gordon's (1988) argument that the interpretation of

contextually-dependent acoustic cues can be based in coarse characteristics of the acoustic signal.

Author's Note

The research reported in this paper was supported by grant #87-305 from the AFOSR Life Sciences Directorate. The author would like to thank Wendy Huey, Shelia Kennison, Christine Nakatani and Annie Senghas for assistance in conducting the research.

References

Cole, R.A., & Rudnicky, A.I. (1983). What's new in speech perception? The research and ideas of William Chandler Bagley, 1874-1946. ***Psychological Review, 90,*** 94-101.

Cutler, A., & Norris, D. (1988). The role of strong syllables in segmentation for lexical access. ***Journal of Experimental Psychology: Human Perception & Performance, 14 (1),*** 113-121.

Denes, P. (1955). Effect of duration on the perception of voicing. ***Journal of the Acoustical Society of America,*** 27, 761-764.

Derr, M.A. & Masssaro, D.W. (1980) The contribution of vowel duration, F0 contour, and friction duration as cues to the /juz/-/jus/ distinction. ***Perception & Psychophysics, 27,*** 51-59.

Gordon, P.C. (1988). Induction of rate-dependent processing by coarse-grained aspects of speech. ***Perception & Psychophysics, 43,*** 137-146.

Grosjean, F. (1980). Spoken word recognition processes and the gating paradigm. ***Perception and Psychophysics, 28 (4),*** 267-283.

Grosjean, F.G., & Gee, J.P. (1987). Prosodic structure and spoken word recognition. In U.H. Frauenfelder & L.K. Tyler (Eds.), ***Spoken Word Recognition*** (pp. 135-155). Cambridge, MA: MIT Press.

Hogan, J.T., & Rozsypal, A.J. (1980). Evaluation of vowel duration as a cue for the voicing distinction of the following word-final consonant. ***Journal of the Acoustical Society of America, 67,*** 1764-1771.

House, A.S., & Fairbanks, G. (1953). The influence of consonant environment upon the secondary acoustic characteristics of vowels. ***Journal of the Acoustical Society of America, 25,*** 105-113.

Klatt, D.H. (1976). Linguistic use of segmental duration in English: Acoustic and perceptual evidence. ***Journal of the Acoustical Society of America, 59,*** 1208-1220.

Ladefoged, P., & Broadbent, D.E. (1957). Information conveyed by vowels. ***Journal of the Acoustical Society of America, 29,*** 98-104.

Lehiste, I., & Peterson, G.E. (1960). Duration of syllable nuclei in English. ***Journal of the Acoustical Society of America, 32,*** 693-703.

Liberman, A.M., Cooper, F.S., Shankweiler, D.P., & Studdert-Kennedy, M.G. (1967). Perception of the speech code. ***Psychological Review, 74,*** 431-461.

Luce, P.A., & Charles-Luce, J. (1985). Contextual effects on vowel duration, closure duration, and and the consonant/vowel ratio in speech production. ***Journal of the Acoustical Society of America, 78,*** 1949-1957.

Martin, J.G. (1970). On judging pauses in Spontaneous speech. ***Journal of Verbal Learning and Verbal Behavior, 9,*** 75-78.

Miller, J.L. (1981). Effects of speaking rate on segmental distinctions. In P.D. Eimas & J.L. Miller (Eds.), ***Perspectives on the study of speech*** (pp. 39-74). Hillsdale, NJ: Erlbaum.

Miller, J.L. (in press). Rate-dependent processsing in speech perception. In A. Ellis (Ed.), ***Progress in the psychology of language*** (Vol. 3). Hillsdale, NJ: Erlbaum.

Miller, J.L., & Baer, T. (1983). Some effects of speaking rate on the production of /b/ and /w/. ***Journal of the Acoustical Society of America, 73,*** 1751-1755.

Miller, J.L., Green, K., & Schermer, T.M. (1984). A distinction between the effects of sentential speaking rate and semantic congruity on word identification. ***Perception and Psychophysics, 36,*** 329-337.

Peterson, G.E., & Barney, H.L. (1952). Control methods used in a study of the vowels. ***Journal of the Acoustical Society of America, 24,*** 175-184.

Pollack, I., & Pickett, J.M. (1963). The intelligibility of excerpts from conversational speech. ***Language and Speech, 6,*** 165-171.

Port, R.F. (1979). The influence of tempo on stop closure duration as a cue for voicing and place. ***Journal of Phonetics, 7,*** 45-56.

Samuel, A.G. (1981). Phonemic restoration: Insights from a new methodology. ***Journal of Experimental Psychology: General, 110 (4),*** 474-494.

Soli, S.D. (1982). Structure and duration of vowels together specify fricative voicing. ***Journal of the Acoustical Society of America, 72 (2),*** 366-378.

Strange, W., Verbrugge, R.R., Shankweiler, D.P., and Edman, T.R. (1976). Consonant environment specifies vowel identity. ***Journal of the Acoustical Society of America, 60,*** 213-224.

Summerfield, Q. (1981). On articulatory rate and perceptual constancy in phonetic perception. ***Journal of Experimental Psychology: Human Perception & Performance, 7,*** 1074-1095.

Umeda, N. (1977). Consonant duration in American English. ***Journal of the Acoustical Society of America, 61,*** 846-858.

Verbrugge, R.R., Strange, W., Shankweiler, D.P., and Edman, T.R. (1976). What information enables a listener to map a talker's vowel space? ***Journal of the Acoustical Society of America, 60,*** 198-212.

Zue, V.W. (1985). Notes on spectogram reading. Unpublished manuscript.

Table 1

The carrier sentences in which the test syllables were embedded. Only sentences 1 and 2 were analyzed and used in the experiments.

1. If Ted says _____ today, Tom will leave the room.
2. If Ted says _____, Tom will leave the room.
3. When Mark reads _____ aloud, Elaine will make a checkmark.
4. When Mark reads _____, Elaine will make a checkmark.

Table 2

The test syllables read by the subjects.

/blz/	/baz/	/biz/	/wlz/	/waz/	/wiz/
/bls/	/bas/	/bis/	/wls/	/was/	/wis/

Table 3

The means and standard deviations (in parentheses) of several durational characteristics of syllable-final /s/ and /z/, broken down by phrasal position, initial consonant identity, and vowel identity. The absolute durational measures are given in msec. The relational measure of proportion transition, presented for syllables beginning with /w/, consists of the duration of the initial second formant transition divided by the that duration plus the duration of the steady state of the vowel.

		/b/			/w/		
		/I/	/a/	/i/	/I/	/a/	/i/
Vowel Duration							
Phrase Internal	/z/	162 (31)	256 (84)	178 (25)	90 (22)	166 (23)	119 (24)
	/s/	118 (25)	191 (13)	140 (20)	52 (14)	130 (7)	69 (12)
Phrase Final	/z/	257 (53)	345 (56)	271 (41)	166 (36)	243 (22)	203 (35)
	/s/	189 (20)	255 (24)	198 (32)	116 (24)	195 (27)	114 (19)
Vowel Offset Duration							
Phrase Internal	/z/	32.3 (14.3)	18.2 (9.5)	29.0 (7.3)	33.7 (11.7)	24.5 (12.5)	25.5 (5.3)
	/s/	14.5 (4.6)	11.2 (4.5)	18.2 (4.3)	17.3 (4.9)	9.3 (2.9)	16.7 (2.7)
Phrase Final	/z/	50.5 (19.0)	33.2 (14.8)	61.8 (24.2)	41.8 (13.0)	40.0 (18.9)	41.3 (18.3)
	/s/	14.5 (4.1)	8.0 (2.3)	18.2 (8.6)	17.5 (4.0)	13.3 (5.0)	14.8 (4.6)

Fricative Duration

Phrase Internal	/z/	79 (35)	95 (45)	88 (42)	103 (21)	96 (28)	90 (42)
	/s/	121 (45)	107 (32)	118 (30)	124 (33)	117 (38)	117 (36)
Phrase Final	/z/	91 (33)	100 (30)	105 (33)	87 (40)	95 (40)	99 (33)
	/s/	150 (37)	163 (70)	157 (81)	142 (50)	158 (71)	181 (71)

Fricative/Vowel Ratio

Phrase Internal	/z/	.50 (.23)	.38 (.18)	.50 (.21)	1.22 (.39)	.58 (.19)	.76 (.34)
	/s/	1.05 (.42)	.57 (.19)	.85 (.23)	2.42 (.46)	.90 (.30)	1.68 (.30)
Phrase Final	/z/	.37 (.16)	.29 (.10)	.41 (.18)	.59 (.36)	.40 (.19)	.50 (.20)
	/s/	.81 (.26)	.66 (.32)	.86 (.53)	1.32 (.64)	.85 (.44)	1.69 (.87)

Proportion Transition

Phrase Internal	/z/				.37 (.08)	.35 (.08)	.29 (.06)
	/s/				.50 (.15)	.34 (.09)	.40 (.05)
Phrase Final	/z/				.30 (.06)	.26 (.03)	.21 (.05)
	/s/				.35 (.06)	.26 (.10)	.29 (.05)

Table 4

Results of Experiment 1. The mean percentage correct identifications of /s/ and /z/ with and without context broken down phrasal position and by initial consonant and vowel.

			Phrase Internal		Phrase Final	
			/s/	/z/	/s/	/z/
Average		Gated	89.8	80.3	74.3	93.5
		In Context	88.9	87.4	80.6	96.5
/b/						
	/I/	Gated	90.3	76.4	79.2	97.2
		In Context	93.1	94.4	77.8	97.2
	/a/	Gated	86.1	81.9	72.2	93.1
		In Context	87.5	93.1	83.3	97.2
	/i/	Gated	86.1	86.1	66.7	87.5
		In Context	81.9	83.3	73.6	88.9
/w/						
	/I/	Gated	95.8	80.6	80.6	97.2
		In Context	94.4	94.4	80.6	100.0
	/a/	Gated	91.7	76.4	76.4	93.1
		In Context	88.9	81.9	77.8	94.4
	/i/	Gated	88.9	80.6	70.8	91.7
		In Context	87.5	80.6	90.3	100.0

Table 5

Results of Experiment 2. The mean percentage correct identifications of /s/ and /z/ with and without context broken down phrasal position and by initial consonant and vowel.

			Phrase Internal		Phrase Final	
			/s/	/z/	/s/	/z/
Average		Gated	90.7	66.9	77.2	90.6
		In Context	92.0	80.4	81.0	91.8
/b/						
	/I/	Gated	94.4	41.7	77.8	90.3
		In Context	97.2	76.4	83.1	90.3
	/a/	Gated	80.6	77.8	74.7	95.8
		In Context	80.6	88.9	73.2	95.8
	/i/	Gated	87.5	65.3	72.2	70.8
		In Context	87.3	73.6	77.8	75.0
/w/						
	/I/	Gated	98.6	62.0	86.1	95.8
		In Context	98.6	90.3	84.7	95.8
	/a/	Gated	93.1	75.0	69.4	97.2
		In Context	93.0	80.6	79.2	94.4
	/i/	Gated	90.3	79.2	83.3	91.7
		In Context	95.8	73.6	87.5	98.6