Duration of frication noise required for identification of English fricatives

Allard Jongman[a)]
Department of Linguistics, Box 1978, Brown University, Providence, Rhode Island 02912

(Received 4 March 1986; accepted for publication 18 October 1988)

Natural speech consonant–vowel (CV) syllables ([f, s, θ, š, v, z, ð] followed by [i, u, a]) were computer edited to include 20–70 ms of their frication noise in 10-ms steps as measured from their onset, as well as the entire frication noise. These stimuli, and the entire syllables, were presented to 12 subjects for consonant identification. Results show that the listener does not require the entire fricative–vowel syllable in order to correctly perceive a fricative. The required frication duration depends on the particular fricative, ranging from approximately 30 ms for [š, z] to 50 ms for [f, s, v], while [θ, ð] are identified with reasonable accuracy in only the full frication and syllable conditions. Analysis in terms of the linguistic features of voicing, place, and manner of articulation revealed that fricative identification in terms of place of articulation is much more affected by a decrease in frication duration than identification in terms of voicing and manner of articulation.

PACS numbers: 43.71.Es

INTRODUCTION

Studies investigating the perception of fricatives have usually focused on various characteristics of the fricative noise portion. In particular, researchers have examined the spectral properties of the noise (Hughes and Halle, 1956; Strevens, 1960), its overall intensity relative to the following vowel (McCasland, 1979), and its duration (You, 1979).

Fricatives are produced with a very narrow constriction in the oral cavity. There is a rapid flow of air through the constriction, creating turbulence in the flow. The random velocity fluctuations in the flow can act as a source of sound. The sound generated in this way is called turbulence noise (Stevens, 1971). Air turbulence produced in this way, by various kinds of constrictions in the vocal tract (the position of which depends on the particular fricative), is the typical sound source for all fricatives.

English fricatives are usually grouped into three or four classes, according to their place of articulation, i.e., according to where the constriction is located in the oral cavity. The overall spectrum shape of the fricative is determined by the size and shape of the oral cavity in front of the constriction, each place of articulation being associated with a distinct spectral shape. However, the location of the formant frequencies in the frication noise varies to some extent from speaker to speaker (Hughes and Halle, 1956) and changes depending on the vowel following the fricative (Soli, 1981).

With regard to duration of frication noise, it seems that different fricatives have different intrinsic durations, although there is some disagreement on which fricatives are longer. Nartey (1982) concluded on the basis of data from several (unspecified) languages that [s, f] are longer than any other fricative. You (1979) showed that, in English, the duration of the frication noise varies with place of articulation: The frication noise of alveolo-palatal fricatives averaged 176 ms, that of alveolars 155 ms, that of labials 103 ms, and that of dentals 99 ms. In addition, Manrique and Massone (1981) found that voiceless Spanish fricatives are longer than voiced ones, a finding that, in general, holds for English as well (Baum and Blumstein, 1987).

In perceptual studies, Harris (1958) and Heinz and Stevens (1961) found that the spectral properties of the frication noise form a major perceptual cue for distinguishing [s] from [š], but that vowel transitions are the primary cue in distinguishing [f, v] from [θ,ð]: perceptual tests showed that the listener uses the vocalic portion to decide which of the sounds [f, v] or [θ,ð] was heard. LaRiviere *et al.* (1975) presented subjects the frication noise in its entirety and obtained high fricative recognition scores for [s, š], somewhat lower scores for [f], and poor scores for [θ]. However, unlike the results of Harris (1958) or Heinz and Stevens (1961), they did not obtain a significant increase in overall recognition scores for [f] or [θ] when adding the vocalic transition to the frication noise. These results are consistent with the observation by Delattre *et al.* (1963) that the vocalic transition does not carry important information for the perception of [f, θ] or their voiced counterparts [v,ð]. LaRiviere *et al.* (1975) suggest, instead, that it is the vowel, rather than the transitions, that appears to play an important perceptual role; subjects obtained significantly higher identification scores for the voiceless fricatives [fa, fu, θi] on the basis of the frication plus the vowel (without the transition), than on the basis of the frication plus the transition. However, as LaRiviere *et al.* point out, the results for [θa] are puzzling: adding either the transition or the vowel to the frication decreased recognition scores for [θa].

Although several studies have looked at properties of the frication noise in its entirety, little is known about the duration of frication noise needed for correct fricative identification. Hughes and Halle (1956) obtained high recognition scores for [s, š, f] using a frication duration of only 50 ms. This study suggests that, at least for these fricatives, listeners do not require the frication noise in its entirety for

[a)] Current address: Central Institute for the Deaf, 818 South Euclid, St. Louis, MO 63110 and Max Planck Institute for Psycholinguistics, P.O. Box 310, NL-6500 AH Nijmegen, The Netherlands.

 0001-4966/89/041718-08$00.80

fricative identification. The present study, then, attempts to determine the minimal duration of frication noise required for correct identification of fricatives.

In order to gain insight into various factors contributing to fricative identification, the data collected were subjected to an analysis in terms of the linguistic features of place of articulation, manner of articulation, and voicing. If, as Hughes and Halle's (1956) results suggested for some voiceless fricatives, duration is not an important perceptual cue for fricative identification, then, given some minimal duration, recognition scores for fricatives should be relatively independent of the frication duration presented. The fact that duration does not play a very important role in distinguishing fricatives in terms of place of articulation would be in accordance with the claim that, for place of articulation, it is the spectral rather than the temporal properties of the acoustic signal that play a role in consonant perception. Stevens and Blumstein (1978), for instance, showed that place of articulation for stop consonants can be accurately identified on the basis of the spectral information that is present in the first 25 ms after the onset of the release burst. Indeed, for fricatives, despite the variability between speakers and the coarticulation effects across vowels, it seems that [s, š, z] can be clearly described and distinguished on the basis of the spectral characteristics of the frication noise alone (Strevens, 1960; Klatt, 1986; Behrens and Blumstein, 1988a). For the other fricatives, notably [f, v] and [θ,ð], it is not clear whether the frication noise, the vowel transition, or some combination of the two would result in a useful characterization. Klatt (1986) and Behrens and Blumstein (1988a) report that the frication noise for both [f] and [θ] has a rather uniform distribution of energy. Strevens (1960) also does not mention any great spectral differences; frequency peaks occur up to and above 7 kHz for both fricatives. The only apparent difference seems to be that [θ] is somewhat lower in intensity than [f] (Klatt, 1986; Behrens and Blumstein, 1988a).

It is unlikely, however, that a duration of about 25 ms would be sufficient for the identification of place and manner of articulation of fricatives, due to the difference in the way the spectral energy builds up in stops and fricatives. Stop consonants are characterized by a rapid spectrum change and simultaneous increase in spectral energy (Pickett, 1980; Stevens, 1980). It might be the case, then, that a longer duration is required by the auditory system for the identification of fricatives, since for a fricative the spectral energy requires more time to reach its maximum amplitude. In fact, in a pilot study (Jongman, 1984) subjects obtained equally high recognition scores for [f, s, š, v, z] on the basis of either 100% or 50% of the frication noise as measured from the onset of frication, but obtained very poor scores on the basis of 10% of the frication noise. This suggests that the cue(s) for place and manner of articulation, as well as voicing, lies somewhere between 10% and 50% of the frication duration, which corresponds to averages of about 14 and 72 ms, respectively.

The purpose of the present study is to establish just how much frication noise is required for correct identification of both voiceless and voiced fricatives and to analyze these data in terms of the linguistic features of place of articulation, manner of articulation, and voicing.

I. METHODS

A. Stimuli

The recorded syllables were consonant–vowel (CV) utterances consisting of the fricatives [f, v, s, z, š, θ, ð] followed by one of the vowels [i, u, a]. Each syllable was printed on a 3-×5-in. notecard in orthographic form. These syllables were produced by one male native speaker of American English.[1] The speaker read all syllables five times in random order at a normal speaking rate. The syllables were recorded on magnetic tape in a soundproof room with a Nagra 4.2 tape recorder and a Shure SM81 microphone. Stimuli were transferred to a PDP 11/34 computer for editing. A 20-kHz sampling rate with a 9-kHz low-pass filter setting was used. For the present experiment, one out of the five repetitions of each target syllable was selected using the following criteria: sound quality, that is, is the target syllable a representative token, as judged informally by three trained phoneticians; and segmentability, that is, those representative tokens were selected for which the onset and offset of the frication noise could be most accurately located.

In the identification task, eight different conditions were presented: the fricative–vowel syllables in their entirety, the frication noise in its entirety (as measured from the onset of the frication noise up to, but not including, the vowel transition), and 70, 60, 50, 40, 30, and 20 ms of the frication noise as measured from the onset of frication. Although the duration of the frication noise of some fricatives is intrinsically longer than that of others (You, 1979; Behrens and Blumstein, 1988a), the frication noise was cut back in absolute steps of 10 ms, rather than in relative steps (as a percentage of the total frication duration). A pilot study (Jongman, 1984) had shown equally good fricative identification scores on the basis of an average of either 144 or 72 ms of frication noise for fricatives whose full frication noise varied in length. Since subjects obtained high recognition scores on the basis of 72 ms, it was decided to cut back the frication noise to 70 ms as the third-longest frication condition (the two longer conditions being the fricative–vowel syllable and the full frication noise).

Measurements were taken directly from a graphics display terminal. First, the onset of the fricative was identified. For [f, θ], this point was sometimes hard to locate, since these fricatives are characterized by a very weak onset of energy. In these cases, the amplitude of the frication noise was scaled up to facilitate visual detection of the noise-onset. Second, the offset of the fricative was identified. For the voiceless fricatives, this point was located at the intensity minimum immediately preceding the onset of periodicity. For the voiced fricatives, the earliest pitch period exhibiting a change in the waveform from that seen throughout the initial frication was identified. This pitch period represented the start of the formant transition into the subsequent vowel. The zero crossing of the preceding pitch period was then designated as the end of the voiced fricative (Yeni-Komshian and Soli, 1981). The durations of the frication noise in

TABLE I. Duration (in ms) of the full frication noise of the test stimuli, for each fricative for each vowel. The numbers in the column on the far right express the percentage of the average frication duration that 70-ms constitutes.

Fricative	Frication duration (ms) Vowel context [a]	[i]	[u]	$\bar{X}$	%
[f]	158	156	134	149	47
[s]	185	193	187	188	37
[θ]	111	88	122	107	65
[š]	153	159	186	166	42
[v]	103	117	118	113	62
[z]	146	165	146	152	46
[ð]	132	82	144	119	59

its entirety are shown in Table I and are comparable to those of Baum and Blumstein (1987).

Five test tapes were made. Ten repetitions of every stimulus were recorded, yielding 7 fricatives $\times 3$ vowels $\times 10$ repetitions $= 210$ stimuli per length condition. The repetitions were randomized and grouped in the following way:

tape 1: the fricative–vowel stimuli in their entirety (210 stimuli);

tape 2: the frication noise in its entirety (210 stimuli);

tape 3: 70, 60, and 50 ms of the frication noise[2] ($3 \times 210 = 630$ stimuli);

tape 4: 40 and 30 ms of the frication noise[2] ($2 \times 210 = 420$ stimuli);

tape 5: 20 ms of the frication noise (210 stimuli).

B. Subjects and procedures

Fourteen male students at Brown University with no formal training in linguistics served as paid subjects. They had no known history of either speech or hearing disorders. The stimuli were presented binaurally through AKG141 headphones at a comfortable level. Subjects responded by writing down one of the 11 alternatives: f, v, s, z, sh, th, tt, p, b, t, d, which were provided at the top of each answering sheet. The alternatives sh, th, and tt were used to indicate [š, θ, ð], respectively. Subjects were asked to repeat a few words with [θ] and [ð] in initial position to ensure that they were aware of the difference between these two sounds. The stop response category was included to provide subjects with an alternative if the fricatives did not sound very fricativelike, a situation that was expected for the short noise conditions. Subjects were told that the number of stop consonants presented was randomized across subjects, so that they might or might not hear any stops in any given length condition. There was a 3-s interval between stimuli and an additional 3 s after every tenth stimulus. To ensure that all subjects understood the instructions, each test tape was preceded by 42 practice stimuli, which were not scored. Subjects were presented not more than one tape per day and finished participating in the experiment in 5 to 8 days. Subjects took a break in the middle of tape 3. The first tape presented was always tape 1, with the fricative–vowel syllables in their entirety, which served as a control task; the order in which the other tapes were presented was randomized for each subject.

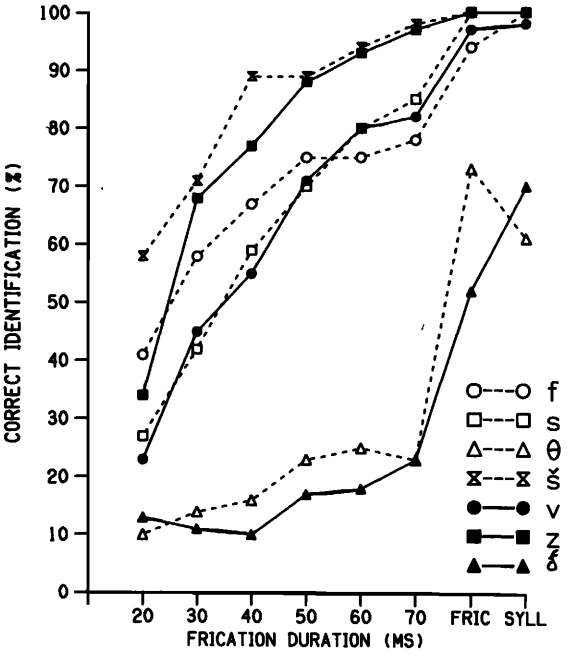

FIG. 1. Correct identification scores (in percent) for each fricative, for each length condition. The scores are averaged over the 12 subjects. "Fric" stands for full frication, "syll" stands for the fricative–vowel syllable.

On the control task, a score of 80% correct was taken as an arbitrary criterion for allowing subjects to participate in the other test conditions. Two subjects did not reach this criterion and did not participate in the remainder of the experiment. The data used in the present experiment were thus obtained from 12 subjects, yielding 12 (subjects) $\times$ 1680 (stimuli) = 20 160 responses.

II. ANALYSIS OF RESULTS

The results section is organized in the following way: First, the correct identification scores for each fricative for each length condition are presented. The next three sections involve analyses in terms of the linguistic features of place of articulation, voicing, and manner of articulation. In addition, fricative identification scores are presented as a function of vowel context. Finally, the data are subjected to information transmission analysis.

A. Identification of fricatives by length condition

Figure 1 shows the mean identification scores for each fricative for each length condition. A two-way ANOVA was performed in which the different fricatives and the different length conditions were compared. There was a main effect both for duration [$F(7,77) = 60.05; p < 0.01$], and for fricative [$F(6,66) = 40.08; p < 0.001$]. For duration, Newman–Keuls *post-hoc* tests revealed that all possible comparisons were significant ($p = 0.01$), except for the difference between the syllable and full frication conditions, that between the 70-, 60-, and 50-ms conditions, and that between the 40- and 30-ms conditions.

For fricatives, Newman–Keuls *post-hoc* tests were also performed to see if there were any significant differences in identification scores for the different fricatives, that is, to see

if any fricative was more accurately identified than any other. In general, highest scores were obtained for [š, z], lower scores for [f, s, v], whereas the lowest scores were obtained for both [θ,ð]. Within each of these groups, there were no statistical differences, except for the difference between [š] and [z] in the 20-ms condition.

There was also a significant fricative × duration interaction effect [$F(42,462) = 6.41$; $p < 0.01$]. Not surprisingly, there is a general trend for fricative identification to increase as the duration of the frication noise gets longer. Newman–Keuls *post-hoc* tests indicated that fricative identification rapidly increased as frication duration increased, reaching a stable value of 70%–80% in the 50-ms condition, and again increasing to around 100% in the full frication and syllable conditions. However, identification scores vary depending on the particular fricative. The fricatives [š, z] were identified with 70% accuracy in the 30-ms condition, whereas similar identification scores for [f, s, v] required a frication duration of 50 ms. For [θ,ð], an increase in frication duration from 20 to 70 ms resulted in only a nonsignificant increase in identification scores; presenting the full frication noise and the entire syllable caused a significant increase in identification, but scores remained well below those for the other fricatives.

In addition to these analyses, separate analyses were performed for correct identification of fricatives in terms of place of articulation, voicing, and manner of articulation.

B. Place of articulation analysis

Table II shows the correct identification scores, averaged over the twelve subjects, for fricative identification in terms of place of articulation, for each length condition, for each fricative.

The scores are presented regardless of identification of voicing and manner of articulation. Following Miller and Nicely (1955), consonants were assigned to one of three categories, depending on the location of the constriction in the vocal tract: "front" ([f, v, p, b]), "middle" ([s, z, θ, ð, t, d]), and "back" ([š]). It is apparent that as the duration of the frication noise increases, the correct identification scores for place of articulation also increase. A two-way ANOVA was performed in which the different fricatives and length conditions were compared. There was a main effect both for duration [$F(7,77) = 57.51$; $p < 0.001$], and for fricative [$F(6,66) = 15.32$; $p < 0.001$]. There was also a significant fricative × duration interaction [$F(42,462) = 4.59$; $p < 0.001$].

Comparing the identification scores for each fricative within each length condition, the differences among fricatives in correct identification of place of articulation gradually decrease as the duration of the frication noise increases, until in the 50-ms condition (and all subsequent longer conditions) identification scores are equal for all fricatives, with the exception of [θ, ð], whose identification scores remain significantly lower than those of the other fricatives.

C. Voicing category

Table III shows the correct fricative identification scores, in terms of voicing, averaged over the twelve subjects, regardless of identification of place and manner of articulation. For this analysis [f, s, θ, š, p, t] were labeled voiceless, and [v, z, ð, b, d] voiced. Overall, fricative identification scores in terms of voicing increase as frication duration increases. A two-way ANOVA was performed in which the different fricatives and the different length conditions were compared. There was a main effect both for duration [$F(7,77) = 9.49$; $p < 0.001$], and for fricative [$F(6,66) = 15.05$; $p < 0.001$]. There was also a significant fricative × duration interaction [$F(42,462) = 5.71$; $p < 0.001$].

Overall, perception of voicing is very stable across all length conditions, with an average percentage of 83% in the 20-ms condition, ranging from 64% for [v] to 96% for [θ]. A comparison of the identification scores for each fricative within each length condition, using Newman–Keuls *post-hoc* tests, reveals that the perception of the three voiced fricatives in the 20-ms condition is significantly poorer that that of the voiceless fricatives, while by the 50-ms condition equal scores for voiced and voiceless fricatives are obtained.

D. Manner of articulation analysis

Table IV shows the correct fricative identification scores in terms of manner of articulation, averaged over the twelve subjects, regardless of identification of place of articulation, and voicing.

TABLE II. Correct fricative identification scores (in percent) in terms of place of articulation, for each fricative, for each length condition. "Fric" stands for full frication noise, "syll" stands for the fricative–vowel syllable.

Frication duration ms	Fricative							
	[f]	[s]	[θ]	[š]	[v]	[z]	[ð]	$\bar{X}$
20	45	61	63	58	48	64	33	53
30	61	69	57	71	64	79	30	62
40	69	71	62	89	68	84	25	67
50	81	84	56	89	83	94	25	73
60	81	86	60	94	84	97	24	75
70	82	88	62	98	86	99	30	78
Fric	99	100	83	100	97	100	68	92
Syll	100	100	74	100	98	100	86	94

TABLE III. Correct fricative identification scores (in percent) in terms of voicing, for each fricative, for each length condition. "Fric" stands for full frication noise, "syll" stands for the fricative–vowel syllable.

Frication duration ms	Fricative							
	[f]	[s]	[θ]	[š]	[v]	[z]	[ð]	$\bar{X}$
20	93	92	96	91	64	68	76	83
30	95	96	98	97	83	95	89	93
40	96	98	97	99	88	99	92	96
50	98	99	99	98	90	100	95	97
60	98	100	99	99	96	99	98	98
70	98	99	99	100	97	100	98	99
Fric	96	100	94	100	99	100	85	96
Syll	100	100	88	99	100	100	84	96

The manner categories were fricative ([f, s, θ, š, v, z, ð]) and stop ([p, t, b, d]). The reader should be cautious in interpreting these results, since there were no stop consonants in the stimulus set. A two-way ANOVA was performed in which the different length conditions and the different fricatives were compared. There was a main effect both for duration [$F(7,77) = 5.04$; $p < 0.0001$], and for fricative [$F(6,66) = 3.41$; $p < 0.001$]. There was also a significant interaction for fricative × duration [$F(42,462) = 2.15$; $p < 0.0001$]. Newman–Keuls *post-hoc* tests were then performed. For most fricatives, there is a significant difference between the 20- and 30-ms conditions on the one hand, and all other (longer) conditions on the other. Comparing the correct identification scores for each fricative within each length condition, it is apparent that identification of manner of articulation increases as the noise duration increases. In the 40-ms condition, and all longer conditions, identification scores for all fricatives but [θ, ð] are statistically equal. Identification scores for [θ] and [ð] become statistically equal to those of the other fricatives in the full frication condition and in the 50-ms condition, respectively.

It is important to note that, although identification of manner of articulation increases as the frication duration increases, identification scores for manner are already very high in the shortest (20 ms) condition, ranging from 87% for [š] to 72% for [ð].

E. Vowel context effect

Table V shows the correct identification scores for the fricatives, for each vowel context, across all length conditions. In order to see if identification of fricatives was influenced by the following vowel, a two-way ANOVA was performed in which the different fricatives and the different vowels averaged over all length conditions were compared. There was a main effect both for fricative [$F(6, 66) = 40.51$; $p < 0.001$], and for vowel [$F(2,22) = 16.33$; $p < 0.001$]. There was also a significant fricative × vowel interaction [$F(12,132) = 2.74$; $p < 0.003$]. Newman–Keuls *post-hoc* tests revealed that subjects identified most fricatives equally well regardless of which vowel followed the frication noise. However, [θ, z,ð] constituted an exception in that these fricatives were generally more often recognized in the context of [i, u], with an average identification score of 51%, than in the context of [a], with an average identification score of 38%.

F. Information transmission analysis

In order to examine any response bias in the identification scores presented above, the data were subjected to an information transmission analysis, as described by Miller and Nicely (1955). Specifically, relative information transmitted was computed for each length condition and for the consonantal features of place of articulation and voicing. Briefly, the confusion matrices for each length condition were collapsed across the features under consideration (e.g., for voicing, voiced, and voiceless consonants were used as stimuli, and voiced and voiceless consonants were used as responses). Total information was then expressed as

$$T(x, y) = -\sum_{i,j} p_{ij} \log_2 \frac{p_i p_j}{p_{ij}},$$

where $T(x, y)$ is transmission in bits per stimulus from x (input) to y (output), p_i is the probability of an input stimulus i, p_j is the probability of an output response j, and p_{ij} is the probability of the joint occurrence of input i and output j. The maximum possible information in the stimulus set, or entropy of the input, is given by

$$H(x) = -\sum_i p_i \log_2 p_i .$$

The values obtained from these equations were then used to calculate the relative information transmitted by

$$T_{\text{rel}}(x, y) = T(x, y)/H(x) .$$

Figure 2 shows the relative information transmitted for each length condition. Information transmitted increases rapidly as frication duration increases, and reaches the 50% level at a duration of approximately 40 ms. It can also be seen that there is a substantial increase in the amount of information transmitted in the full frication condition as compared to the 70 ms condition.

Figure 3 shows the relative information transmitted for fricative identification in terms of place of articulation for each length condition. In general, information transmitted for place increases as frication duration increases. At a frication duration of approximately 70 ms, 50% of the total information is transmitted. Comparing this analysis to the results

TABLE IV. Correct fricative identification scores (in percent) in terms of manner of articulation, for each fricative, for each length condition. "Fric" stands for full frication noise, "syll" stands for the fricative–vowel syllable.

Frication duration ms	Fricative							
	[f]	[s]	[θ]	[š]	[v]	[z]	[ð]	$\bar{X}$
20	82	85	81	87	79	79	72	81
30	88	84	77	93	84	88	77	84
40	93	98	79	99	92	92	81	91
50	93	89	76	98	93	97	93	91
60	93	94	81	99	97	99	95	94
70	95	96	85	100	97	100	95	95
Fric	99	100	96	100	100	100	100	99
Syll	100	100	99	100	100	100	99	100

TABLE V. Correct fricative identification scores (in percent), for each fricative, for each vowel context.

Fricative	Vowel context			
	[i]	[a]	[u]	$\bar{X}$
[f]	75	72	74	74
[s]	70	70	71	70
[θ]	32	24	38	31
[š]	87	87	87	87
[v]	69	67	71	69
[z]	87	71	89	82
[ð]	29	20	33	27
$\bar{X}$	64	59	63	

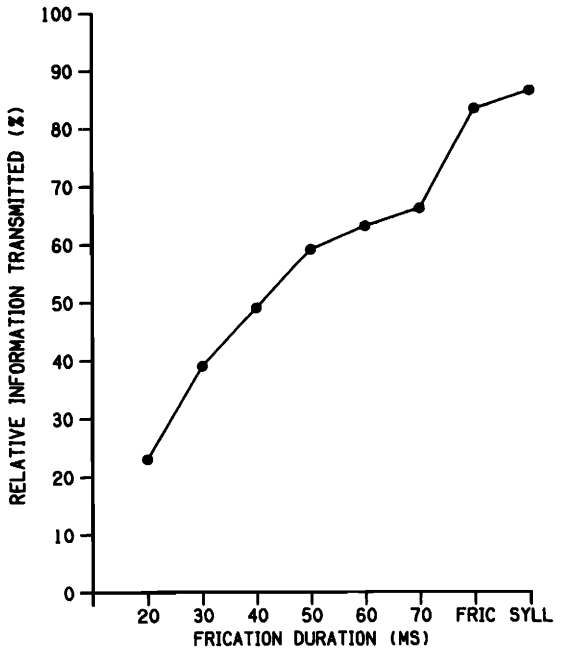

FIG. 2. Relative information transmitted (in percent) for each length condition. "Fric" stands for full frication, "syll" stands for the fricative–vowel syllable.

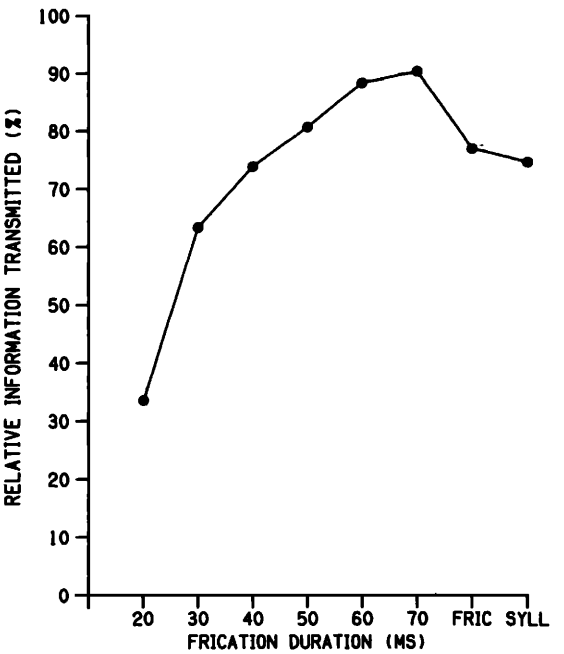

FIG. 4. Relative information transmitted (in percent) for fricative voicing, for each length condition. "Fric" stands for full frication, "syll" stands for the fricative–vowel syllable.

in Table II, we see that both analyses show an increase in identification with an increase in frication duration, with the full frication condition yielding substantially better identification in terms of place of articulation.

Figure 4 shows the relative information transmitted for fricative identification in terms of voicing for each length condition. The information transmitted for voicing reaches a high level around a frication duration of 30 ms. Comparing this figure to Fig. 3, it is obvious that identification of voicing is much less affected by a decrease in frication duration than identification of place of articulation.

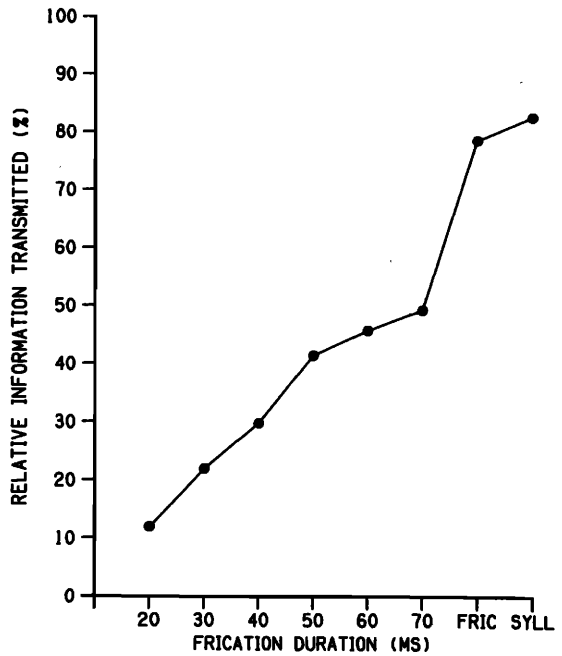

FIG. 3. Relative information transmitted (in percent) for fricative place of articulation, for each length condition. "Fric" stands for full frication, "syll" stands for the fricative–vowel syllable.

III. DISCUSSION

The present experiment was designed to determine the amount of frication noise that is required by the listener in order to correctly identify the fricatives of American English. Overall, and perhaps not surprisingly, identification of fricatives improves as the duration of the frication noise increases. Information transmission analysis shows that, at a frication duration of 40 ms, 50% of the total information is transmitted. However, it can be said that the minimum frication duration required for identification varies from fricative to fricative. For [š, z], approximately the initial 30 ms of the frication noise enables listeners to identify these fricatives with good accuracy, whereas 50 ms or more is required for [f, s, v]. Identification scores for [θ, ð], however, only reach a reasonably high level in the full frication and syllable conditions.

Perception of fricatives in terms of place of articulation was more affected by a decrease in the duration of the frication noise than perception of fricatives in terms of voicing and manner of articulation. Identification of place of articulation gradually increased as the frication duration increased, with more than 50% of the total information for place being transmitted in only the two longest conditions.

Recent research on place of articulation in stop consonants suggests that there are stable acoustic cues that specify place of articulation in initial stop–vowel syllables (Stevens and Blumstein, 1978; Blumstein and Stevens, 1979; Kewley-Port, 1983; Lahiri *et al.*, 1984). Although these researchers have somewhat different views on what constitutes a stable

(invariant) cue and on where in the signal it resides, both Blumstein and Stevens and Kewley-Port support the claim that sufficient properties for identification of place of articulation are found in the initial 20 to 40 ms of a natural stop–vowel syllable: Perception experiments have shown that listeners can correctly identify place of articulation in stops on the basis of the burst and the first few tens of milliseconds of formant transitions (Blumstein and Stevens, 1980; Kewley-Port *et al.*, 1983).

The present results, however, indicate that the first 20–40 ms of the frication noise of a fricative–vowel syllable do not contain sufficient information for identification of place of articulation in fricatives, with the exception of [š, z]. The fact that a longer duration is required for identification of place of articulation in fricatives, as compared to that in stops, might be due to the gradual increase in amplitude for fricatives versus the abrupt onset of energy in stops. If, as Stevens and Blumstein (1978) claim, it is the case that the spectral characteristics of the acoustic signal carry the major information for place of articulation in stop consonants, then 20–40 ms of a stop–vowel syllable should be sufficient for identification of place of articulation in stops, since the spectrum change for stops is complete in 10–30 ms (Stevens, 1980). Fricatives, on the other hand, are characterized by a less rapid spectral change, which might explain why the listener needs more than 40 ms in order to be able to identify place of articulation in fricatives.

Identification scores for fricatives in terms of voicing were high for all length conditions. The present results do not support the findings by Cole and Cooper (1975), who found that the duration of the frication noise is a sufficient cue for distinguishing among the voiced and voiceless fricatives. In their study, the syllables [fa] and [sa] were shortened by removing segments of frication noise just prior to the vowel, after which the initial portion of the fricative was spliced back onto the vowel. In their perception experiment, Cole and Cooper found that shortening the duration of the frication noise produced a marked shift in the percept from voiceless to voiced. However, the present results show that subjects do not have a tendency to identify more fricatives as voiced as the frication duration decreases. Identification of the voicing category is quite good, independent of the duration of the frication noise, and frication noise duration *per se* can, therefore, not be considered a sufficient cue in the distinction among the voiced and voiceless fricatives.

Identification scores for manner of articulation (fricative versus stop) were high, even in the shortest condition, for which the average identification score was approximately 72%. In other words, subjects presented with the first 20 ms of the frication noise identified only 28% of these stimuli as stops. However, at even shorter noise durations, fricatives are predominantly perceived as stops. In a previously mentioned pilot study (Jongman, 1984), subjects identified manner of articulation for [f, s, š, v, z] with only 13% accuracy on the basis of an average frication duration of approximately 14 ms. This suggests, then, that the minimal duration for the perception of fricatives in terms of manner of articulation may lie somewhere between 14 and 20 ms. A potential acoustic cue could be the gradual amplitude increase characteristic of fricative spectra (Stevens, 1980).

Research on the perception of fricatives has mainly focused on two characteristics of the frication noise: spectral characteristics and amplitude characteristics. For example, with respect to spectral properties, it has been suggested that vocalic information (vocalic transitions and/or steady-state vowel) is the primary cue in identifying [f, v, θ, ð], since the noise portion of these fricatives does not carry much information (Harris, 1958; Heinz and Stevens, 1961). Our results, however, show good perception (around 95%) for both [f] and [v] on the basis of the frication noise alone. Furthermore, the fact that there is no significant difference in the identification of either the fricatives [θ, ð] or [f, v] in the full frication condition compared to the syllable condition shows that perception of both [θ, ð] and [f, v] does not improve when the vocalic transitions and steady-state vowel are added to the frication noise.

With respect to amplitude properties, the present results suggest that noise intensity *per se* is unlikely to be a major cue to fricative identification. Although the rank order in fricative identification for all length conditions [ranging approximately (high to low) from [š, z] to [s, f, v] to [θ, ð]) shows a resemblance to the rank order of intensity of the voiceless fricatives as reported by Strevens (1960) (Strevens did not report on voiced fricatives), amplitude measurements of the full frication noise of the test stimuli used in the present experiment revealed the following order: [z] (70 dB), [v] and [ð] (66 dB), [s] (65 dB), [š] (64 dB), [θ] (54 dB), and [f] (53 dB). This order is similar to that reported by Behrens and Blumstein (1988b) for the voiceless fricatives. Although [s] and [š] are of comparable intensity, identification of [š] is better than that of [s]. Similarly, identification of [f] is better than that of [θ], though they are of comparable intensity. Identification of [ð] with a noise of relatively high intensity is poorer than that of most fricatives with lower intensity. Thus, on the basis of the present results, it seems unlikely that noise intensity *per se* would be a major cue in distinguishing fricatives. One way to study the role of noise amplitude in identification of fricatives would be to conduct a perception experiment in which all fricatives are presented at the same sound pressure level.

Finally, our findings concerning the effects of vowel context on fricative identification are somewhat puzzling. Although for all fricatives (but [θ, z, ð]), identification is not significantly affected by vowel context, for [θ, z, ð] fricative identification tends to be better in the context of [i, u] than in the context of [a]. These results differ from those obtained by Yeni-Komshian and Soli (1981). Presenting 150 ms of frication noise as measured back from vowel onset, Yeni-Komshian and Soli (1981) found that identification of the fricatives [s, z, š, ž] was significantly better in the context of the vowel [a] than in the context of [i, u]. According to Soli (1981), one would expect identification of these fricatives in the [a] context to be the least affected by anticipatory vowel coarticulation, since acoustic analysis of [s, z, š, ž] indicated that fricatives produced in the [a] context had spectral characteristics that were most similar to those of fricatives produced in isolation.

In conclusion, the present research suggests that the

first 50 ms of frication noise, corresponding to less than half of the total noise duration, contains sufficient information for fricative identification, with the exception of [θ, ð]. Thus the listener does not require the frication noise in its entirety in order to correctly perceive a fricative. The critical duration does depend somewhat on the fricative under consideration, ranging from approximately 30 ms for [š, z] to 50 ms for [f, s, v], while [θ] and [ð] are identified with reasonable accuracy in only the full frication and syllable conditions. More research is needed in order to determine if, and to what extent, spectral and temporal properties, as well as amplitude characteristics, of this 50-ms initial segment of the speech signal contribute to correct identification of fricatives.

ACKNOWLEDGMENTS

The author would like to thank Susan Behrens, Sheila Blumstein, Philip Lieberman, Joanne Miller, Joan Sereno, and Sigfrid Soli for helpful discussions and suggestions. This research was supported in part by Grant NS15123 to Brown University.

[1]Although the fricative stimuli of the present study were produced by one speaker, there is strong reason to believe that the results will generalize to tokens of other speakers. Though different speakers may employ different articulatory strategies in fricative production, many studies (e.g., Strevens, 1960; Behrens and Blumstein, 1988a) present data averaged across speakers because they find interspeaker differences to be negligible. In particular, after individual analyses of the spectral, temporal, and amplitude properties of 100 fricative tokens produced by each of three male speakers, Behrens and Blumstein (1988a) concluded that there is "tremendous consistency in the pattern of results found for the three subjects."

[2]Stimuli consisting of 70, 60, and 50 ms of the frication noise were randomized and put together on one tape, since in the pilot study (Jongman, 1984) neither the subjects nor the experimenter perceived them as having very different durations. Stimuli consisting of 40 and 30 ms of the frication noise were grouped together for the same reason.

Baum, S. R., and Blumstein, S. E. (**1987**). "Preliminary observations on the use of duration as a cue to syllable-initial fricative consonant voicing in English," J. Acoust. Soc. Am. **82**, 1073–1077.

Behrens, S. J., and Blumstein, S. E. (**1988a**). "Acoustic characteristics of English voiceless fricatives: A descriptive analysis," J. Phonet. **16**, 295–298.

Behrens, S. J., and Blumstein, S. E. (**1988b**). "On the role of the amplitude of the fricative noise in the perception of place of articulation in voiceless fricative consonants," J. Acoust. Soc. Am. **84**, 861–867.

Blumstein, S. E., and Stevens, K. N. (**1979**). "Acoustic invariance in speech production: Evidence from measurements of the spectral characteristics of stop consonants," J. Acoust. Soc. Am. **66**, 1001–1017.

Blumstein, S. E., and Stevens, K. N. (**1980**). "Perceptual invariance and onset spectra for stop consonants in different vowel environments," J. Acoust. Soc. Am. **67**, 648–662.

Cole, R. A., and Cooper, W. E. (**1975**). "Perception of voicing in English affricates and fricatives," J. Acoust. Soc. Am. **58**, 1280–1287.

Delattre, P. C., Liberman, A. M., and Cooper, F. S. (**1963**). "Formant transitions and loci as acoustic correlates of place of articulation in American English fricatives," Stud. Linguist. **18**, 104–121.

Harris, K. S. (**1958**). "Cues for the discrimination of American English fricatives in spoken syllables," Lang. Speech **1**, 1–7.

Heinz, J. M., and Stevens, K. N. (**1961**). "On the properties of voiceless fricative consonants," J. Acoust. Soc. Am. **33**, 589–596.

Hughes, G. W., and Halle, M. (**1956**). "Spectral properties of fricative consonants," J. Acoust. Soc. Am. **28**, 303–310.

Jongman, A. (**1984**). "On the duration of frication noise as a perceptual cue" (unpublished manuscript).

Kewley-Port, D. (**1983**). "Time-varying features as correlates of place of articulation in stop consonants," J. Acoust. Soc. Am. **73**, 322–335.

Kewley-Port, D., Pisoni, D. B., and Studdert-Kennedy, M. (**1983**). "Perception of static and dynamic cues to place of articulation in initial stop consonants," J. Acoust. Soc. Am. **73**, 1779–1793.

Klatt, D. H. (**1986**). "Voiceless fricatives" (unpublished manuscript).

Lahiri, A., Gewirth, L., and Blumstein, S.E. (1984). "A reconsideration of acoustic invariance for place of articulation in diffuse stop cononants: Evidence from a cross-language study," J. Acoust. Soc. Am. **76**, 391–404.

LaRiviere, C., Winitz, H., and Herriman, E. (**1975**). "The distribution of perceptual cues in English prevocalic fricatives," J. Speech Hear. Res. **18**, 613–622.

Manrique, A. M. B., and Massone, M. I. (**1981**). "Acoustic analysis and perception of Spanish fricative consonants," J. Acoust Soc. Am. **69**, 1145–1153.

McCasland G. P. (**1979**). "Noise intensity and spectrum cues for spoken fricatives," J. Acoust. Soc. Am. Suppl. 1 **65**, S78–S79.

Miller, G. A., and Nicely, P. E. (**1955**). "An analysis of perceptual confusions among some English consonants," J. Acoust. Soc. Am. **27**, 338–352.

Nartey, J. N. A. (**1982**). "On fricative phones and phonemes," *UCLA Working Papers in Linguistics*, No. 55.

Pickett, J. M. (**1980**). *The Sounds of Speech Communication* (University Park, Baltimore).

Soli, S. D. (**1981**). "Second formants in fricatives: Acoustic consequences of fricative–vowel coarticulation," J. Acoust. Soc. Am. **70**, 976–984.

Stevens, K. N. (**1971**). "Airflow and turbulence noise for fricative and stop consonants: Static considerations," J. Acoust. Soc. Am. **50**, 1182–1192.

Stevens, K. N. (**1980**). "Acoustic correlates of some phonetic categories," J. Acoust. Soc. Am. **68**, 836–842.

Stevens, K. N., and Blumstein, S. E. (**1978**). "Invariant cues for place of articulation in stop consonants," J. Acoust. Soc. Am. **64**, 1358–1368.

Strevens, P. (**1960**). "Spectra of fricative noise in human speech," Lang. Speech **3**, 32–49.

Yeni-Komshian, B., and Soli, S. D. (**1981**). "Recognition of vowels from information in fricatives: Perceptual evidence of fricative–vowel coarticulation," J. Acoust. Soc. Am. **70**, 966–975.

You, H.-Y. (**1979**). "An acoustical and perceptual study of English fricatives," M. A. thesis, University of Edmonton, Edmonton, Canada.