# The relative perceptual distinctiveness of initial and final consonants in CVC syllables

Melissa A. Redford[a] and Randy L. Diehl
*Department of Psychology, University of Texas, Austin, Texas 78712*

(Received 24 November 1997; revised 3 February 1999; accepted 3 May 1999)

Among the world's languages, syllable inventories allowing only initial consonants predominate over those allowing both initial and final consonants. Final consonants may be disfavored because they are less easy to identify and/or more difficult to produce than initial consonants. In this study, two perceptual confusion experiments were conducted in which subjects identified naturally produced consonant–vowel–consonant syllables in different frame sentences. Results indicated that initial consonants were significantly more identifiable than final consonants across all conditions. Acoustic analyses of the test syllables indicated that the relative identifiability of initial and final consonants might be explained in terms of production differences as indicated by the greater acoustic distinctiveness of initial consonants. 
[S0001-4966(99)04009-6]

PACS numbers: 43.71.Es [WS]

## INTRODUCTION

It is natural to assume that sound patterns that are widely attested among languages may be explained in terms of human biology. According to one such class of explanations, preferred sound patterns arise in response to performance factors involving perception, production, or both (e.g., Lindblom *et al.*, 1984; Diehl and Lindblom, 1996). This class of explanations has been used to account for sound patterns in infant babbling (Davis and MacNeilage, 1995), prevalent feature cooccurrences within sounds in languages (Diehl and Kluender, 1989), and the structure of vowel systems (Lindblom, 1986) and consonantal systems (Willerman, 1994).

Bell and Hooper (1978) made a variety of observations regarding cross-language syllable structure with reference to word units. For instance, they noted that (1) initial consonants predominate over final consonants and that final vowels predominate over initial vowels, (2) the consonants that occur in final position are generally a subset of the consonants that can occur in initial position, and (3) consonant clusters are disfavored across languages, but are more likely to occur in initial position rather than final position. These typological facts suggest that syllable inventories across languages favor syllables with simple (single) consonantal onsets over those with consonantal offsets.

It is also likely that syllables with initial consonants predominate even in languages that allow syllable-final consonants. For instance, the most frequent syllable type in English is the consonant–vowel (CV) syllable (Greenberg, 1998). Historical processes also indicate a preference for initial consonants over final consonants. For example, a common sound change is final consonant weakening and loss (Hock, 1986). In this paper, we examine whether the syllable structure preference for initial consonants over final consonants may be attributed to perceptual and/or production factors. We first report results of two perceptual confusion experiments designed to assess the relative identifiability of initial and final consonants, and then provide evidence, in the form of acoustic measurements on the test stimuli, that the perceptual results may be explained in part by production differences between initial and final consonants.

Ample evidence suggests that syllable-initial and syllable-final consonants are perceived differently by listeners. The selective adaptation paradigm, for example, has shown that repetitive presentation of CV syllables fails to elicit adaptation effects in the perception of vowel–consonant (VC) syllables composed of the same vowel and consonant (Ades, 1974; Wolf, 1978; Samuel, 1989). However, evidence regarding a perceptual advantage for syllables with consonantal onsets over those with consonantal offsets is relatively scant and inconclusive because the question has not been the focus of research.

Van Wieringen (1995) reviewed studies of plosive identification as a function of formant transitions into or out of the syllabic nucleus and reported that eight studies found that the initial (CV) transition was perceptually more distinctive, while the remaining six studies found that the final (VC) transition was more distinctive. Evidence based on other acoustic cues also leads to conflicting conclusions regarding the relative identifiability of initial and final consonants. For instance, Raphael and Dorman (1980) found that the shorter the interval of silence between a final consonant and a following initial consonant, the more subjects report hearing only the initial consonant. On the other hand, Winitz *et al.* (1972) found that voiceless stop release bursts of similar duration with and without accompanying vowel formant transitions provide at least as much information about consonant place of articulation for phrase-final as for phrase-initial consonants.

It should be noted that Raphael and Dorman (1980),

[a]Electronic mail: mredford@utxsvs.cc.utexas.edu

Winitz *et al.* (1972), and each of the studies reviewed by Van Wieringen (1995) used either synthetic or digitally altered stimuli which were typically presented in isolation. As such, the studies were not designed to address the question of whether naturally produced consonants are generally more identifiable in one syllable position versus another.

Results of perceptual confusion studies that use naturally produced, unaltered stimuli have also been somewhat mixed, although they tend to show a perceptual advantage for initial consonants over final ones. Miller and Nicely (1955) found that, overall, initial stop consonants in isolated CV syllables (V=/a/) were less confusable than final stops in isolated consonant–vowel–consonant (CVC) syllables (with CV =/ta/). Wang and Bilger (1973) reported that consonants were less confusable in CV than in VC syllables when the vowel was /a/ or /u/, but the reverse was true when the vowel was /i/. Unlike Miller and Nicely (1955) or Wang and Bilger (1973), Helfer and Huntley (1991) embedded CV and VC test syllables in a carrier phrase, thus approximating more natural listening conditions. In many cases they found no position differences in consonant identifiability, but when such differences did occur, initial consonants were favored over final consonants.

While previous studies lend some support to the claim that initial consonants are more identifiable than final ones, a more complete test is needed. First, it is appropriate that the target syllables occur within a carrier phrase (as in Helfer and Huntley, 1991) so as to approximate more natural conditions of phonetic production and perception than those found in previous perceptual confusion experiments. Constraints that apply only to the production and perception of isolated syllables would appear to be unlikely sources of cross-language syllable structure preferences. Second, once a decision is made to embed the target syllables in a carrier phrase, a question arises about how best to structure the phonetic context in the immediate vicinity of each syllable. With the exception of /s/, English consonants in initial position occur invariably before a vowel or sonorant consonant. Final stop consonants often occur before a syllable-initial consonant and therefore might lack place cues associated with a vocalic or sonorant environment. Accordingly, a general test of the relative identifiability of consonants across initial and final syllable position should include opportunities for the final consonant to appear before both a vowel and a consonant. Finally, if consonant identifiability turns out to be position dependent, acoustic analyses should be performed to evaluate whether this perceptual effect may be related to position-dependent differences in consonant production.

We therefore designed a perceptual confusion experiment to directly test the relative identifiability of initial and final consonants under conditions that could be generalized to the conditions under which phonological systems emerge. Naturally spoken CVC syllables were presented in frame sentences, which provided both a consonantal and a vocalic context for the final consonant.

The segments /p, t, k, f, θ, s, ʃ/ and /i, u, a/ were used to create the CVC syllables. With the exception of /θ/, these segments have a wide distribution in the world's languages. According to Maddieson's (1984) analyses of 317 languages from diverse language families, the plain voiceless stops /p, t, k/ are the most commonly occurring consonants in languages (92% of the languages examined have all three), /s/ is the most common voiceless fricative in languages which have fricatives (84%), followed by /ʃ/ and /f/ (46% and 43%). The vowels /i/, /u/, and /a/ are by far the most common among world's languages (92%, 84%, and 88%, respectively).

Predictions were based on the hypothesis that syllable structure preferences across languages reflect differences between initial and final consonants in perceptual distinctiveness. Specifically, it was predicted that initial consonants would yield fewer overall perceptual confusions errors than final consonants. We also predicted that this position effect would occur across consonants, vowel nuclei, and talkers. This assumes that differences in perceptual distinctiveness are associated with syllabic position generally and not with phoneme- or talker-specific production contexts and conditions. A final prediction was based on the typological fact that clusters are disfavored, but more likely to occur in initial position rather than in final position. Because preconsonantal final consonants might behave as though they formed part of a consonant cluster, it was predicted that they would be more confusable than prevocalic final consonants.

## I. PERCEPTUAL EXPERIMENT 1

### A. Method

#### *1. Stimuli*

The consonants /p, t, k, f, θ, s, ʃ/ and the vowels /i, a, u/ were combined to form CVC target syllables. Each consonant occurred with every other consonant in both syllable positions and with three different vowel nuclei ($7\times7\times3=147$).[1] Two male and two female speakers of American English produced each target syllable in two frame sentences ($147\times4\times2=1176$). The frame sentences which yielded a consonantal context for the syllable-final consonant was *Say* CVC *some more*. The vocalic context was produced with *Say* CVC *again*. The sentences were recorded using the Kay Elemetrics Computerized Speech Laboratory with a Shure SM48 microphone. The root-mean-squared (rms) amplitude of each stimulus sentence was normalized across talkers using a waveform editor developed in our laboratory. The rms amplitude was calculated for the entire utterance, which was then scaled so that the average amplitude of the entire utterance was consistent across all utterances.

#### *2. Subjects and procedure*

Subjects were seven female and seven male native American-English-speaking college students. They were instructed to listen for the CVC target syllables embedded in the two frame sentences and to write each syllable in normal orthography on the response sheet provided. Subjects were informed what consonants and vowels would appear in the syllable set, but were encouraged to write down whatever they heard even if it did not correspond to these sounds. To simulate natural listening conditions, stimulus sentences were presented in a low level of pink noise[2] (15-dB signal-to-noise ratio). Subjects were seated in a second-attenuated

TABLE I. Confusion matrices for initial consonants (top) and for preconsonantal (middle) and prevocalic (bottom) final consonants. Subjects' response types are listed horizontally at the top of each matrix. The "other" and "none" columns refer to responses outside of the token set and null responses, respectively.

| | p | t | k | f | θ | s | ʃ | Other | None | |
|---|---|---|---|---|---|---|---|---|---|---|
| p | **2295** | 15 | 23 | 9 | | 1 | | | 9 | |
| t | 6 | **2254** | 5 | 24 | 3 | 1 | 1 | 5 | 53 | T |
| k | 39 | 14 | **2251** | 4 | 6 | 2 | | 2 | 34 | a |
| f | 384 | 24 | 20 | **1678** | 113 | 65 | | 14 | 54 | r |
| θ | 129 | 44 | 12 | 670 | **1274** | 197 | 3 | 7 | 16 | g |
| s | 10 | 25 | 1 | 84 | 87 | **2106** | 27 | 3 | 9 | e |
| ʃ | 1 | | 2 | 2 | 3 | 38 | **2302** | 2 | 2 | t |

| | p | t | k | f | θ | s | ʃ | Other | None | |
|---|---|---|---|---|---|---|---|---|---|---|
| p | **1007** | 37 | 38 | 32 | 42 | 2 | | 15 | 3 | |
| t | 76 | **950** | 38 | 11 | 72 | 2 | | 20 | 7 | T |
| k | 32 | 13 | **1063** | 15 | 14 | 1 | | 30 | 8 | a |
| f | 431 | 35 | 90 | **366** | 184 | 19 | 6 | 37 | 8 | r |
| θ | 249 | 52 | 104 | 122 | **589** | 23 | | 23 | 14 | g |
| s | 17 | 29 | 9 | 41 | 180 | **748** | 81 | 30 | 41 | e |
| ʃ | 1 | | | 3 | 7 | 31 | **1119** | 11 | 4 | t |

| | p | t | k | f | θ | s | ʃ | Other | None | |
|---|---|---|---|---|---|---|---|---|---|---|
| p | **1005** | 4 | 3 | 73 | 60 | 7 | | 18 | 6 | |
| t | 25 | **1090** | 7 | 3 | 24 | 1 | | 16 | 10 | T |
| k | 27 | 10 | **1084** | 9 | 11 | 1 | | 26 | 8 | a |
| f | 192 | 15 | 23 | **660** | 218 | 28 | 3 | 31 | 6 | r |
| θ | 53 | 29 | 18 | 184 | **810** | 54 | 1 | 22 | 5 | g |
| s | 11 | 10 | 5 | 34 | 90 | **911** | 84 | 23 | 8 | e |
| ʃ | | | | | 1 | 23 | **1131** | 19 | 2 | t |

room and listened to the stimuli over earphones. The sentences occurred in a randomized sequence at interstimulus intervals of 3.5 s.

## B. Results

There were a total of 2352 observations for each consonant in each syllable position. Perceptual confusions in manner or place (but not voicing[3]) were recorded as errors, as were nonresponses. Confusion matrices are presented in Table I for all initial consonants and for final consonants in a consonantal context and final consonants in a vocalic context. Visual inspection of the matrices reveals that the overall pattern of confusions was broadly similar between the initial and the two final positions: in all three cases, the error rates were lowest for /ʃ/, somewhat higher for the stop consonants /p/, /t/, and /k/, higher still for /s/, and highest for the nonsibilant fricatives /f/ and /θ/.

The relative percentage of errors were subjected to an arcsine transform to ensure a normal distribution, and a repeated measures analysis of variance (ANOVA) was performed with four within-subjects factors (syllable position/context, target consonant, vowel nucleus, and talker). The syllable position factor consisted of three levels: initial, final before a consonant, and final before a vowel. There were significant main effects of syllable position [$F(2,26) = 38.18$, $p<0.001$], target consonant [$F(6,78) = 52.68$, $p<0.001$], and talker [$F(3,39) = 142.18$, $p<0.001$], but not of vowel nucleus [$F(2,26) = 1.73$, $p>0.1$].

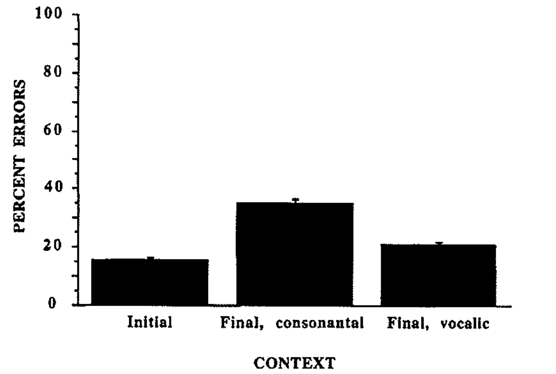


FIG. 1. Identification error rate on initial and preconsonantal and prevocalic final consonants in experiment 1.

Figure 1 displays consonant identification error rates in initial position, final position before a consonant, and final position before a vowel. As predicted, subjects made more errors identifying final consonants than initial consonants irrespective of context. Although final consonants were more identifiable when preceding a vowel, they were still less identifiable than the initial consonants. Planned comparisons revealed significant differences between initial position and final position before a consonant [$F(1,26) = 21.96$, $p<0.001$], between initial position and final position before a vowel [$F(1,26) = 5.61$, $p<0.05$], and between the two final position contexts [$F(1,26) = 37.21$, $p<0.001$].

There were significant interactions between syllable position and consonant [$F(12,156) = 17.48$, $p<0.001$] and syllable position and speaker [$F(6,78) = 5.07$, $p<0.001$]. Mean comparisons statistically confirm what visual inspection of Fig. 2 (top panel) indicates, namely, that the perceptual advantage associated with initial position over final consonants in either context was observed for all four talkers. The position effect occurred for all consonants except /θ/ (Fig. 2, middle panel). Mean comparisons indicated, however, that apart from /θ/, two of the seven consonants (/k/ and /ʃ/) were not statistically more identifiable in initial position than in preconsonantal final position. In addition, three of the seven consonants (/t/, /k/, and /ʃ/) were as identifiable in prevocalic final position as in initial position. It is possible that failure to achieve statistical significance in these cases resulted from insufficient power (i.e., not enough errors). The interdental fricative /θ/ was the only consonant that reversed the general finding that initial consonants were more perceptible than final consonants in any context. Prevocalic final /θ/ was correctly identified significantly more often than initial /θ/ or preconsonantal final /θ/. No significant differences in identifiability occurred, however, between initial /θ/ and preconsonantal /θ/.

There was also a significant interaction between syllable position and vowel nucleus [$F(4,52) = 5.46$, $p<0.001$] (Fig. 2, bottom panel). Although mean comparisons indicated that preconsonantal final consonants were statistically less identifiable than initial consonants across all vowel nuclei, final consonants that preceded a vowel were only statistically less

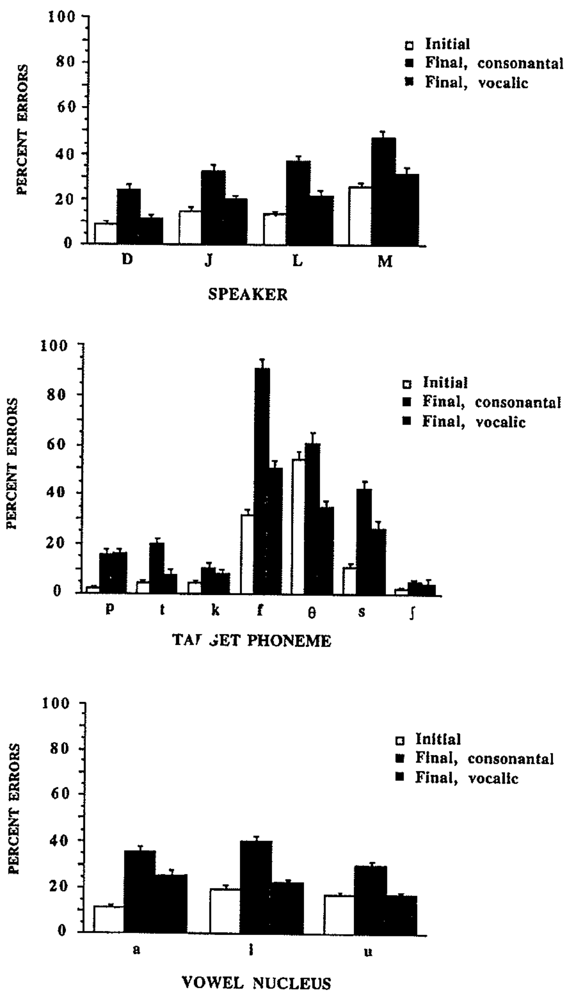


FIG. 2. Identification error rates on initial and final consonants across the different talkers (top), different phonemes (middle), and different vowel nuclei (bottom) of experiment 1.

identifiable than initial consonants when the vowel nucleus was /a/.

### C. Discussion

Results of this experiment provide support for the prediction that, in naturally produced CVC syllables, initial consonants are less confusable than final consonants. The prediction that this position effect would occur across consonants, vowel nuclei, and talkers was also generally upheld. Even though the effect did not reach significance in every case, it was reversed only in the case of /θ/. The final prediction—that final consonants that occur before a consonant would be less identifiable than those that occur before a vowel—was also supported, albeit not as strongly. In the overall analysis, final consonants in a consonantal context were less identifiable than those in a vocalic context, but a significant interaction with syllable position and vowel nucleus suggested that this effect might depend on the preceding vowel context.

The fact that prevocalic final consonants that follow /i/ or /u/ nuclei were found to be as identifiable as initial consonants also points to a potential problem in this experiment: the perceptual advantage observed here for consonants in initial position could instead be a perceptual advantage for intervocalic consonants. This alternative interpretation of the main result is supported by evidence that suggests that vowel transitions both into and out of a vowel carry information about the adjacent consonant (Winitz *et al.*, 1972; Ostreicher and Sharf, 1976). This possibility is not necessarily contradicted by the fact that prevocalic final consonants are less identifiable than initial consonants when the vowel nucleus is /a/ because the pre- and postvocalic environments of the initial and final consonants are not identical. In initial position, as in final position, the prevocalic environment is always a stressed vowel—in initial position this vowel is the /e/ of "say" and in final position it is one of the three vowel nuclei. The postvocalic environment, however, differs in the two conditions. For initial consonants the postvocalic environment is also always a stressed vowel, but for final consonants the following vowel is always the unstressed, reduced vowel /ə/ of "again." The discrepancy between the pre- and postvocalic environments of the final consonants may be especially important when the vowel nucleus is /a/ given that this vowel is generally produced with greater amplitude than other vowels. The greater discrepancy between the environments might explain why prevocalic final consonants are less perceptible than initial consonants when the vowel nucleus is /a/.

In light of the possibility that the main result from experiment 1 may reflect greater identifiability of intervocalic consonants rather than of initial consonants *per se*, we conducted another perceptual confusion experiment with the same CVC tokens embedded in different frame sentences. To test whether the main result of experiment 1 indeed reflected a perceptual advantage for initial consonants over final consonants, experiment 2 placed both initial and final consonants in a consonantal context as well as in a stressed, vocalic context.

## II. PERCEPTUAL EXPERIMENT 2

### A. Method

#### *1. Stimuli*

As in the first experiment, the consonants /p, t, k, f, θ, s, ʃ/ and the vowels /i, a, u/ were combined to form CVC target syllables. Each consonant occurred with every other consonant in both syllable positions and with three different vowel nuclei ($7\times7\times3=147$). Two male and two female speakers of American English produced each target syllable in two frame sentences ($147\times4\times2=1176$). Only one of the speakers was the same as in experiment 1. The frame sentences differed from experiment 1. The sentence which yielded a consonantal context for the target syllable was *I said* CVC *definitively*, which provided a similar consonantal context for both the initial and final consonants of the target syllable. The vocalic context was produced with *Say* CVC *eight times*, which ensured that the vocalic context was as similar as possible for the initial and final consonants of the target syllable. As in experiment 1, the sentences were recorded using the Kay Elemetrics Computerized Speech Laboratory with a

TABLE II. (a) Confusion matrices for postconsonantal (top) and postvocalic (bottom) initial constants. Subjects' response types are listed horizontally at the top of each matrix. The 'other' and 'none' columns refer to responses outside of the token set and null responses respectively. (b) Confusion matrices for preconsonantal (top) and prevocalic (bottom) final consonants. Subjects' response types are listed horizontally at the top of each matrix. The "other" and "none" columns refer to responses outside of the token set and null responses respectively.

(a)

| | p | t | k | f | θ | s | ʃ | Other | None | |
|---|---|---|---|---|---|---|---|---|---|---|
| p | **1109** | 11 | 17 | 19 | 1 | 1 | | | 18 | |
| t | 45 | **1082** | 12 | 5 | 7 | 4 | 1 | 1 | 19 | T |
| k | 25 | 17 | **1108** | 3 | | | | 4 | 19 | a |
| f | 87 | 16 | 7 | **981** | 23 | 22 | | 1 | 39 | r |
| θ | 25 | 44 | 9 | 435 | **340** | 257 | 10 | 2 | 54 | g |
| s | 4 | 59 | 6 | 45 | 21 | **971** | 47 | 2 | 21 | e |
| ʃ | 1 | | 1 | 3 | 2 | 72 | **1071** | 7 | 19 | t |

| | p | t | k | f | θ | s | ʃ | Other | None | |
|---|---|---|---|---|---|---|---|---|---|---|
| p | **1105** | 17 | 11 | 23 | 1 | 1 | | | 18 | |
| t | 36 | **1075** | 35 | 6 | 1 | | 1 | | 22 | T |
| k | 31 | 15 | **1090** | 11 | 1 | 1 | 1 | 3 | 23 | a |
| f | 113 | 13 | 20 | **941** | 28 | 20 | | 1 | 40 | r |
| θ | 24 | 41 | 18 | 456 | **340** | 241 | 7 | 3 | 46 | g |
| s | 2 | 17 | 4 | 68 | 29 | **981** | 49 | 3 | 23 | e |
| ʃ | 1 | 2 | 1 | 1 | 1 | 63 | **1087** | 6 | 14 | t |

(b)

| | p | t | k | f | θ | s | ʃ | Other | None | |
|---|---|---|---|---|---|---|---|---|---|---|
| p | **709** | 153 | 172 | 29 | 53 | 5 | | 11 | 44 | |
| t | 117 | **875** | 88 | 8 | 34 | 4 | 1 | 4 | 45 | T |
| k | 152 | 96 | **750** | 26 | 69 | 6 | 1 | 11 | 65 | a |
| f | 174 | 56 | 199 | **322** | 228 | 10 | 10 | 12 | 165 | r |
| θ | 118 | 111 | 122 | 141 | **424** | 73 | 10 | 16 | 161 | g |
| s | 19 | 24 | 38 | 25 | 61 | **794** | 54 | 95 | 66 | e |
| ʃ | | 5 | | 3 | 11 | 40 | **1059** | 40 | 18 | t |

| | p | t | k | f | θ | s | ʃ | Other | None | |
|---|---|---|---|---|---|---|---|---|---|---|
| p | **720** | 102 | 139 | 49 | 71 | 17 | 1 | 5 | 72 | |
| t | 90 | **830** | 65 | 17 | 61 | 12 | 1 | 1 | 99 | T |
| k | 35 | 48 | **1019** | 9 | 9 | 4 | 1 | 2 | 49 | a |
| f | 162 | 43 | 128 | **292** | 281 | 66 | 13 | 14 | 177 | r |
| θ | 107 | 89 | 118 | 154 | **428** | 90 | 8 | 13 | 169 | g |
| s | 11 | 16 | 44 | 12 | 45 | **853** | 84 | 37 | 74 | e |
| ʃ | | | 1 | 1 | | 27 | **1088** | 39 | 20 | t |

Shure SM48 microphone. The root-mean-squared (rms) amplitude of each stimulus sentence was normalized across talkers. The rms amplitude was calculated for the last half of the frame sentence (e.g., for *definitively* and *eight times*). The entire utterance was then scaled so that the average amplitude of the latter part of the frame sentence would be consistent across all utterances. In this way the stimuli were normalized across talkers, but the inherent amplitude differences of the different CVC target syllables were preserved even better than in experiment 1.

### 2. Subjects and procedure

Subjects were seven female and seven male native American-English-speaking college students, none of whom had served in the previous experiment. The instructions and procedure were identical to those of experiment 1.

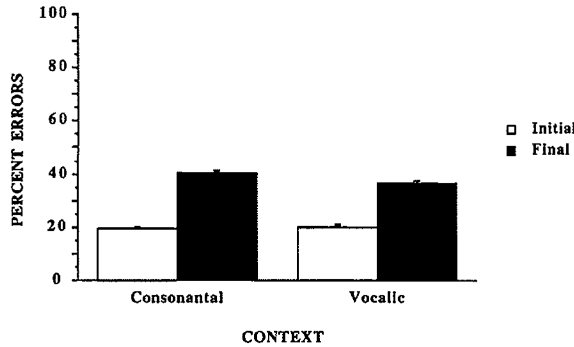


FIG. 3. Identification error rates on initial and preconsonantal and prevocalic final consonants in experiment 2.

### B. Results

There were a total of 2352 observations for each consonant in each syllable position. Perceptual confusions of manner and place (but not voicing) were recorded as errors, as were nonresponses. Confusion matrices for initial and final consonants in the two conditions are presented in Table II. As in experiment 1, the pattern of confusion errors was qualitatively similar across conditions: the fricative /ʃ/ was confused least often of any phoneme, and the stop consonants and /s/ were confused less often than the fricative /f/ and /θ/.

A repeated measures analysis of variance (ANOVA) was performed on error rate with five within-subjects factors (syllable position, context, target consonant, vowel nucleus, and talker). There were significant main effects for syllable position [$F(1,13)=105.01$, $p<0.001$] and context [$F(1,13)=12.41$, $p<0.005$], target consonant [$F(6,78)=74.8$, $p<0.001$], vowel nucleus [$F(2,26)=13.75$, $p<0.001$], and talker [$F(3,39)=174.37$, $p<0.001$].

There was a significant interaction between syllable position and context [$F(1,13)=35.5$, $p<0.001$] (see Fig. 3). Planned comparisons reveal that subjects made more errors identifying final consonants than initial consonants irrespective of context ([$F(1,13)=1534.67$, $p<0.001$] in a consonantal context and [$F(1,13)=945.5$, $p<0.001$] in a vocalic context), and on preconsonantal final consonants than on prevocalic final consonants [$F(1,13)=55.25$, $p<0.001$], but that postconsonantal initial consonants were no less identifiable than postvocalic initial consonants [$F(1,13)=0.99$, $p>0.1$]. In spite of the statistically significant difference between the identifiability of prevocalic and preconsonantal final consonants, the graph shows clearly that this difference is small. Thus, further analyses deal only with the difference between initial and final positions.

Although significant interactions occurred between syllable position and target consonant [$F(6,78)=27.66$, $p<0.001$], syllable position and vowel nucleus [$F(2,26)=11.7$, $p<0.001$], and syllable position and talker [$F(3,39)=16.86$, $p<0.001$], the perceptual advantage associated with initial position was observed for all four talkers (Fig. 4, top panel), every consonant except /θ/ (Fig. 4, middle panel), and all three vowel nuclei (Fig. 4, bottom panel). Mean comparisons indicated that these differences

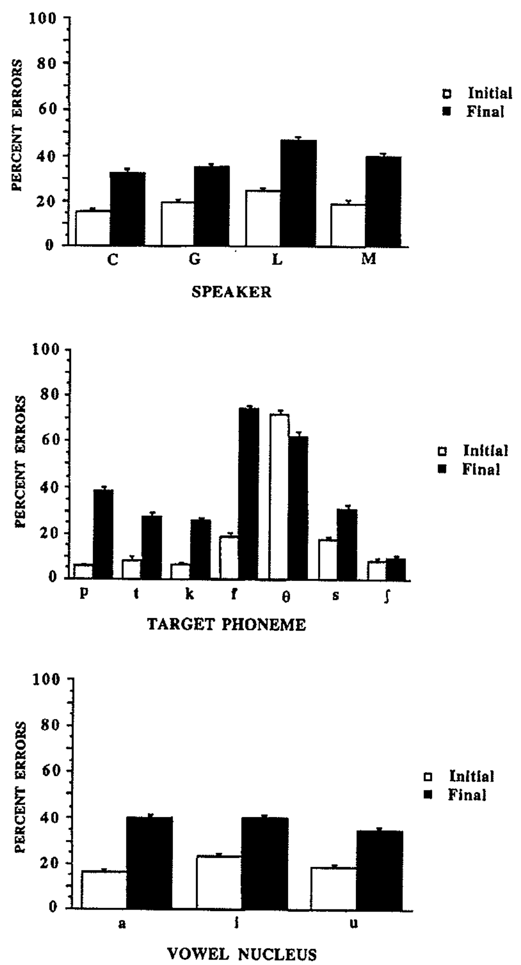


FIG. 4. Identification error rates on initial and final consonants across the different talkers (top), different phonemes (middle), and different vowel nuclei (bottom) of experiment 2.

were statistically significant ($p<0.05$) in all conditions for all cases except /ʃ/, which had the fewest errors of any of the consonants in either position.

## C. Discussion

### 1. Predictions

Results from the second perceptual confusion experiment confirm that the perceptual advantage associated with initial consonants over final consonants is attributable to position rather than to context. The findings from the two experiments are consistent with demonstrations of a perceptual advantage for syllable onsets over syllable offsets (e.g., Raphael and Dorman, 1980) and with results of some earlier perceptual confusion studies (Miller and Nicely, 1955; Helfer and Huntley, 1991). The results do, however, deviate in one respect from those of Wang and Bilger (1973), who reported a perceptual *dis*advantage for initial consonants when the vowel nucleus was /i/. The reasons for this discrepancy between the two studies are unclear. In any case, the observed positional effect on consonant identifiability is consistent with the hypothesis that cross-language preferences for initial over final consonants emerge through selection for communicative value.

Another prediction was that the position effect would occur across consonants, vowel nuclei, and talkers. This was also confirmed. Although in both experiments there were significant interactions between position and each of the other variables, the greater identifiability of initial consonants occurred for six of the seven consonants, all three vowel nuclei, and all four talkers, and thus appears to be quite general. Moreover, the overall pattern of confusions was similar across positions, with sibilant fricatives and stop consonants being more accurately identified than nonsibilant fricatives in both cases.

It is not clear why the position effect was reversed in the case of /θ/. In any case, on the basis of this result, it might be expected that /θ/, already a rare phoneme in the world's languages, will occur preferentially in inventories, like English, that allow for final consonants.

The prediction that syllable-final consonants would be less confusable before a vowel than before another consonant was strongly upheld in experiment 1, but only weakly in experiment 2. This discrepancy between experiments may stem from differences between the vowels that followed the final consonants. In Experiment 1, the final consonant preceded an unstressed vowel. Phonologists have noted that consonants in this context are often treated as ambisyllabic (Kahn, 1976) or as geminates (Burzio, 1994), making the consonant syllable-initial as well as syllable-final. In experiment 2, the final consonant preceded a stressed vowel. Hoard (1966) noted that glottal stops are often inserted before stressed, word-initial vowels. Visual inspection of our stimulus waveforms confirmed that glottal stops were typically inserted in this context. The presence of a glottal stop would preclude resyllabification of the final consonant across the word boundary. Thus, whereas the prevocalic final consonants in experiment 1 may have been treated (also) as syllable-initial consonants, those of experiment 2 were likely to be treated only as syllable-final ones.

### 2. Explanations

There are at least two types of explanation for why initial consonants are more easily identified than final consonants. One explanation attributes the position effect to differences in auditory processing and assumes that similar acoustic information is available for both initial and final consonants. A candidate auditory processing account of this type comes from a study by Delgutte and Kiang (1984), who presented isolated synthetic CV and VC stimuli to anesthetized cats and measured discharge patterns of auditory-nerve fibers. They found that the discharge rate was lower for consonants in syllable-final position because of short-term adaptation to the preceding vowel. On the basis of this evidence, it might be argued that syllable onsets are more readily perceived than syllable offsets because onsets yield a greater neuronal response. However, such an account does not appear to be applicable to consonant perception in phrase contexts, where syllable-initial consonants are also susceptible to the effects of short-term adaptation from preceding vowels or consonants. Thus, the account does not readily explain either the present findings or the cross-language position preferences in syllable structure that motivated this study.

Another possible explanation for the position effect on identification accuracy is that initial consonants may be produced differently than final consonants, resulting in different acoustic correlates. For example, there is some evidence that consonants tend to be longer in syllable-initial than in syllable-final position, provided the syllable is not in phrase-final position (Hoard, 1966; Anderson and Port, 1992). In the case of fricatives, a shorter duration might yield less overall information for segment identity. For stops, a short duration could negatively affect identification in the context of other consonants. Raphael and Dorman (1980) found that shorter silent intervals (corresponding to stop closure) reduced perceptibility of the first stop in a sequence of two consonants straddling a syllable boundary. Also, segments that are produced more quickly tend to be more reduced and more variable (Moon and Lindblom, 1994; Byrd and Tan, 1996), especially in final position (Byrd, 1996).

Another potentially relevant factor in the perception of consonants is amplitude. Sounds that are harder to hear are harder to identify. It is likely that the repeatedly observed difference in identifiability between the sibilant fricatives, /s/ and /ʃ/, and the nonsibilant fricatives, /f/ and /θ/, is largely attributable to their differences in relative amplitude. With respect to stop consonants, Henderson and Repp (1982) noted that although preconsonantal stops usually have release bursts, these are of much lower amplitude than prevocalic stops and therefore difficult to detect. Given that bursts are important (and sometime sufficient) cues for the identification of voiceless stops (Winitz *et al.*, 1972), differences in perceptibility may result if syllable position affects their amplitude (Malecot, 1968).

Winitz *et al.* (1972) also found that the average frequency of the stop release burst provided good place-of-articulation information. Place information is also provided by the average frequency value of the noise in fricative consonants (Harris, 1958). Given that average frequency plays a role in the identification of consonants, positional variation in frequency values may affect the relative perceptibility of the consonants to be identified.

Another perceptual cue for consonants is the vowel formant transitions into and out of the consonant. These transitions reflect the coarticulation of the consonant with the following or preceding vowel and provide information for consonant place of articulation (Delattre *et al.*, 1955; Harris, 1958) and manner (Liberman *et al.*, 1956). In a review of the acoustic and perceptual literature on stop consonants, Pickett *et al.* (1995) conclude that CV transitions are generally more informative about stop consonant identity than VC transitions. Results of studies on "locus equations" are consistent with this conclusion. Locus equations are regression functions describing the relationship between the onset (or offset) frequency of the second formant ($F2$) in a CV (or VC) transition and the $F2$ value of the adjacent vowel nucleus. These variables have been shown to be linearly related for a given place category across different vowels, and different place categories are associated with different locus equation slopes (Lindblom, 1963). Krull (1988) and Sussman *et al.* (1997) found that these slopes differ less across place of articulation categories for VC transitions than for CV transition, indicating that $F2$ trajectories are more distinctive for place value in CV than in VC syllables. In addition, Sussman *et al.* (1997) found that the variance explained by locus equations is significantly greater for stops in initial than in final position. Finally, it is important to note that locus equation parameters have been shown to be relevant for listeners' identification of place categories (Fruchter and Sussman, 1997, but see commentaries associated with Sussman *et al.*, 1998).

In view of the preceding discussion, five different types of acoustic measurements were made on the CVC stimuli used in experiment 2 to determine whether differences in production could account for the results obtained. A discriminant analysis was then performed to determine how accurately consonants could be categorized on the basis of these measurements.

## III. ACOUSTIC MEASUREMENTS

### A. Method

Five acoustic measurements—one of duration, one of amplitude, an average frequency measure, and two frequency measures related to $F2$ transitions—were made on each of the initial and final consonants of the CVC stimuli produced by the four talkers in experiment 2. Only the stimuli from experiment 2 were analyzed because this experiment provided the stronger test of the relative identifiability of initial and final consonants. Fricative duration was measured from the onset to the offset of friction noise. In the case of stops, the durational measurement included the release burst and aspiration interval as well as the closure interval. (In cases where unreleased final stops preceded another stop consonant, duration of the target consonant could not be determined.) Root-mean-squared (rms) amplitude was calculated for energy present during the entire constriction interval for fricatives and for the energy present in a 51.2-ms window centered at the onset of the release burst for stops. A Hamming window was applied to these same intervals, and the frequency centroid [$=\Sigma(x_i y_i)/\Sigma y_i$, where $x_i$= frequency in Hz of $i$th sample and $y_i$= corresponding amplitude], or average frequency, was calculated. The $F2$ transition measures were taken at the onset of $F2$ for initial consonants (or the offset of $F2$ for final consonants) and at the midpoint of the following (or preceding) vowel. These last two measures correspond to the data points used in studies of locus equations. Following Sussman and Shore (1996), the $F2$ onset (or offset) values were measured at the point where $F2$ was first (or last) clearly distinguishable in a visual inspection of the spectrogram.

### B. Results

Table III presents the average values of the five acoustic measurements for the initial and final consonants of the CVC stimuli from experiment 2. The values presented are averaged across 84 observations for each of the fricatives in each condition and for each of the stops in the two initial conditions. The values presented for final stops are averaged

TABLE III. Acoustic measurements of the initial and final consonants of the CVC stimuli from experiment 2. Average values are displayed for each of the phonemes in the two contexts: consonantal (top) and vocalic (bottom). Duration measures are in milliseconds, frequency measures are in Hz, and amplitude measures are rms values calculated from 16-bit integer signals.

| | Duration | | Amplitude | | Average frequency | | $F2$ onset/offset | | $F2$ midpoint | |
|---|---|---|---|---|---|---|---|---|---|---|
| | Initial | Final | Initial | Final | Initial | Final | Initial | Final | Initial | Final |
| p | 153 | 103 | 599 | 270 | 1774 | 1776 | 1664 | 1661 | 1807 | 1859 |
| t | 161 | 111 | 808 | 542 | 2664 | 2381 | 2181 | 1962 | 1953 | 1902 |
| k | 152 | 103 | 634 | 315 | 2304 | 2012 | 2161 | 2012 | 1847 | 1876 |
| f | 136 | 105 | 286 | 381 | 2480 | 2358 | 1673 | 1731 | 1781 | 1824 |
| θ | 135 | 106 | 268 | 270 | 2265 | 2058 | 1933 | 1896 | 1913 | 1904 |
| s | 155 | 115 | 666 | 821 | 3356 | 3365 | 1958 | 1919 | 2155 | 1897 |
| ʃ | 172 | 121 | 1996 | 1813 | 3387 | 3253 | 2112 | 2069 | 1898 | 1868 |
| p | 151 | 122 | 575 | 314 | 1764 | 1545 | 1721 | 1644 | 1801 | 1873 |
| t | 158 | 109 | 706 | 467 | 2653 | 2422 | 2254 | 2032 | 1952 | 1906 |
| k | 159 | 122 | 614 | 550 | 2327 | 2432 | 2196 | 2116 | 1853 | 1874 |
| f | 153 | 133 | 280 | 326 | 2407 | 2324 | 1675 | 1719 | 1795 | 1840 |
| θ | 146 | 136 | 212 | 245 | 2187 | 1989 | 2024 | 1902 | 1904 | 1884 |
| s | 175 | 140 | 630 | 633 | 3342 | 3330 | 2074 | 1956 | 1918 | 1885 |
| ʃ | 178 | 140 | 2042 | 1771 | 3362 | 3294 | 2182 | 2103 | 1910 | 1873 |

across the released final stops only and are based on 63, 33, and 52 observations for preconsonantal /p/, /t/, and /k/, respectively, and 83, 73, and 84 observations for prevocalic /p/, /t/, and /k/, respectively.

Repeated measures analyses of variance (ANOVA) were performed on each of the measurements on initial and final consonants (except for $F2$ of the vowel nucleus) with five factors (syllable position, context, target consonant, vowel nucleus, and talker). The analyses indicated that initial consonants were greater both in duration [$F(1,6)=937.82$, $p<0.001$] and amplitude [$F(1,6)=101.3$, $p<0.001$] than final consonants. Average frequency and frequency of $F2$ onset/offset also differed between initial and final consonants [$F(1,6)=349.2$, $p<0.001$; $F(1,6)=9.4$, $p<0.05$].

A significant interaction between syllable position and target consonant occurred for all measurements. Significant interactions between syllable position and context occurred for duration, amplitude, and average frequency. Significant interactions also occurred for duration and amplitude between syllable position and talker and for average frequency between syllable position and vowel nucleus. Mean comparisons indicated that (1) consonant duration was significantly greater in initial position than in final position across all conditions; (2) consonant amplitude was significantly greater in initial position than in final position for both contexts, all vowel nuclei, and for /p/, /t/, /k/, and /ʃ/, but not for /f/, /θ/, or /s/; (3) average frequency and the onset/offset frequency of $F2$ transition were significantly higher in initial position than final position, but varied considerably across the different conditions. Even though there was no effect for intensity for three of the seven consonants, there were no significant reversals of the tendency for initial consonants to be longer and more intense than final consonants.

Pearson correlations on 168 token-measure pairs were made for each of the consonants in the initial and final conditions. Table IV lists coefficients for correlations between

TABLE IV. Correlation coefficients between number of errors and each of the measures for each of the initial and final consonants.

| | Duration | | Amplitude | | Average frequency | | $F2$ onset/offset | |
|---|---|---|---|---|---|---|---|---|
| | Initial | Final | Initial | Final | Initial | Final | Initial | Final |
| p | 0.133 | 0.251[b] | −0.185[a] | −0.218[b] | 0.091 | −0.060 | 0.039 | 0.220[b] |
| t | −0.050 | −0.192[a] | −0.255[b] | −0.244[b] | −0.180[a] | −0.105 | 0.250[b] | 0.219[b] |
| k | 0.138 | −0.045 | −0.362[b] | −0.484[b] | 0.036 | −0.346[b] | 0.049 | −0.044 |
| f | 0.175[a] | −0.068 | −0.084 | −0.317[b] | −0.055 | −0.203[b] | −0.126 | −0.178[a] |
| θ | 0.084 | −0.043 | −0.063 | −0.237[b] | −0.149 | −0.098 | 0.252[b] | −0.060 |
| s | −0.133 | −0.039 | −0.439[b] | −0.508[b] | −0.543 | −0.618[b] | 0.342[b] | 0.102 |
| ʃ | −0.103 | 0.095 | −0.172[a] | −0.079 | −0.132 | −0.003 | −0.111 | 0.071 |

[a] $p<0.05$.
[b] $p<0.01$.

each of the acoustic measures (excluding $F2$ of the vowel nucleus) and listeners' identification error rate for each consonant. A quite consistent inverse relationship was observed between consonant amplitude and error. However, shorter consonant durations were not associated with more errors.

A discriminant analysis was performed to assess the degree to which the tokens could be correctly classified on the basis of the five acoustic measures. Classification accuracy is presented in Fig. 5. Overall correct classification for initial consonants was 66.0% in the consonantal context and 67.2% in the vocalic context. Correct classification for final consonants was 54.3% in the consonantal context and 58.8% in the vocalic context. The performance patterns are similar to those of the human subjects, but the human subjects outperform discriminant analysis in most cases (see Fig. 4 for comparison with human subjects). Only for /f/ and /θ/ did discriminant analysis outperform human subjects.

Table V shows the standardized canonical discrimination coefficients of the first two discriminant functions. These coefficients reflect the degree to which a given acoustic variable contributes successfully to the classification of stimulus tokens. Across consonant class and position, the most highly weighted coefficients corresponded to the frequency measures, especially $F2$ transition onset or offset and $F2$ of the preceding or following vowel. The average frequency measure and amplitude were somewhat less important than the $F2$ transition measures in discriminating phoneme categories, but more important than duration.

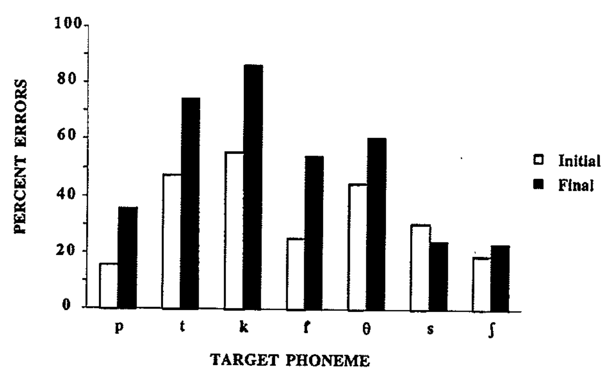


FIG. 5. Correct classification by disciminant analysis of each phoneme in initial and final position.

TABLE V. Standardized canonical discriminant coefficients for the five discriminant functions.

| | Duration | | Amplitude | | Frequency | | $F2$ onset/offset | | $F2$ midpoint | |
|---|---|---|---|---|---|---|---|---|---|---|
| | Initial | Final | Initial | Final | Initial | Final | Initial | Final | Initial | Final |
| 1 | 0.085 | 0.071 | 0.643 | 0.483 | 0.592 | 0.765 | 0.439 | 0.091 | −0.438 | −0.149 |
| 2 | −0.017 | 0.029 | 0.658 | 0.898 | −0.866 | −0.701 | 0.685 | 0.116 | −0.329 | 0.028 |
| 3 | 0.077 | −0.611 | 0.419 | −0.088 | −0.034 | 0.037 | −1.235 | 0.812 | 0.804 | 0.118 |
| 4 | 1.016 | 0.785 | −0.187 | −0.121 | −0.096 | −0.062 | −0.041 | 0.583 | −0.156 | −0.221 |
| 5 | 0.046 | 0.153 | 0.01 | 0.025 | −0.032 | −0.015 | −0.099 | −0.174 | 1.067 | 0.995 |

### C. Discussion

Analyses of the acoustic measurements made on the CVC stimuli suggest that the perceptual advantage associated with initial consonants may have a basis in production. Initial consonants were found to be longer and louder than their final counterparts. Similar position differences in duration and amplitude have been previously reported in the literature (Hoard, 1996; Anderson and Port, 1992; Malecot, 1968; Henderson and Repp, 1982). The link between amplitude and perceptibility is obvious and was supported by the tendency for number of errors to be significantly inversely correlated with amplitude. The link between frequency differences and identifiability might reflect the relative role that frequency plays in differentiating consonant types. For example, the fact that higher frequencies correlated positively with identification accuracy for /s/ indicates that subjects probably used frequency as a cue to distinguish /s/ from the other fricatives. The expected inverse correlation between duration and number of errors was not found.

A discriminant analysis based on acoustic measurements of duration, amplitude, average frequency, and $F2$ transition yielded more accurate classification of initial consonants than final consonants. In this respect, discriminant classification performance was similar to human performance in the identification experiment. The fact that initial consonants were better classified suggests they are produced with greater acoustic distinctiveness than final consonants.

Despite the parallels, humans generally outperformed discriminant analysis except in the classification of nonsibilant fricatives. The advantage for human listeners is not surprising since they have access to a considerably richer array of acoustic cues than are available to a discriminant classifier. For instance, one important cue not available to the classifier was the presence or absence of a silent interval, normally used by human listeners to distinguish stops from fricatives. Also, because the discriminant analysis only included released stops, the relative contribution of a release burst *per se* is unknown.

## IV. GENERAL DISCUSSION

The purpose of this study was to evaluate whether cross-language preferences for syllable-initial consonants may be influenced by perceptual and/or production factors. Consonant identification experiments were conducted to test three predictions: (1) that syllable-initial consonants are identified more accurately than syllable-final consonants; (2) that this perceptual advantage occurs across consonants, vowel nuclei, and talkers; and (3) that syllable-final consonants preceding a vowel are identified more accurately than those preceding a consonant. The first prediction was clearly supported. The second prediction was supported in that the position effect on identification accuracy was observed for all four talkers, all three vowel nuclei, and for six of the seven consonants, albeit not always at statistically significant levels. In addition, the overall pattern of identification errors was similar across initial and final positions in both experiments: sibilant fricatives and stops were more accurately identified than nonsibilant fricatives.

The third prediction—that syllable-final consonants are more identifiable before a vowel than before a consonant—received strong support in experiment 1, but not in experiment 2. As discussed earlier, it is possible that the discrepancy between the experiments arises from a difference in vocalic environment in the two cases. Specifically, in experiment 1 the target final consonant occurred before a schwa—an environment conductive to ambisyllabicity (or gemination)—making the consonant syllable-initial as well as syllable-final. However, in experiment 2, the target final consonant was followed by a stressed word-initial vowel, an environment that typically prompts the insertion of a glottal stop and thus precludes resyllabification of the final consonant.

Acoustic properties of the stimuli were analyzed to determine whether or not production differences could account for the observed differences in identification accuracy. A significant positive correlation between amplitude and accuracy, but not between duration and accuracy, indicated that amplitude contributed more than duration to differences in perceptibility of initial and final consonants. The discriminant analysis indicated that frequency measures, especially those associated with the $F2$ transition, also contributed strongly to the separation of the consonant categories. Importantly, discriminant classification performance was greater for initial than for final consonants, suggesting that the perceptual advantage for initial consonants may be attributable to their greater acoustic distinctiveness.

The present study was motivated by the view that factors of perception and production shape cross-language preferences in sound structure. The result that initial consonants are more identifiable than final consonants is consistent with the typological observation that CV syllables are preferred to syllables with final consonants. The fact that this result held in consonantal as well as vocalic contexts is consistent with the cross-language observation that consonant sequences, though disfavored overall, are more likely to occur in initial position than in final position.

The perceptual advantage associated with initial consonants held across consonant and adjacent vowel categories, indicating that the syllable position effect is not attributable to specific target phonemes and contexts. However, the general nature of the effect does not preclude the possibility that different consonants are more or less suited for final position. This possibility is consistent with the typological observation that consonants permitted in final position represent a small subset of a language's total consonant inventory.

It is, of course, possible that the results of the present study are limited to the language environment in which they were observed. On the other hand, it might be expected that because English syllable structure incorporates many different syllable types with final consonants, English speakers and listeners would be especially adept at the production and perception of final consonants. Nevertheless, the broad cross-language arguments made here would be considerably strengthened by replicating these findings with speakers of other languages.

Although our results indicate a perceptual advantage for syllable-initial consonants, it is possible that the advantage is restricted to word-initial consonants. Further research involving multi-syllabic words is necessary to clarify this issue.

In summary, cross-language preferences for syllable-initial consonants may be attributable to their relatively high identifiability. This perceptual advantage may derive in part from production factors giving rise to greater acoustic distinctiveness.

## ACKNOWLEDGMENTS

This paper benefited from the insightful criticisms and comments of three anonymous reviewers. The work was supported by a NSF Graduate Research Fellowship to the first author and by Research Grant No. 5 R01 DC00427-09, -10 from the National Institute on Deafness and Other Communication Disorders, National Institutes of Health, to the second author.

[1]Unavoidably, some of the combinations corresponded to actual English words. Since lexical status of the syllables is independent of the syllable position and context variables, there is no reason to expect any biasing effect of lexical status with respect to the hypotheses under test.

[2]Pink noise was used because it has a spectrum that drops off in amplitude at the same rate as speech. The flat spectrum of white noise would mask the higher more than the lower frequencies of speech.

[3]Voicing errors were not recorded because the stimuli did not vary along this dimension and subjects were told in advance which consonants and vowels they would hear in the experiment.

Ades, A. E. (**1974**). "How phonetic is selective adaptation? Experiments on syllable position and vowel environment," Percept. Psychophys. **16**, 61–66.

Anderson, S., and Port, R. (**1992**). "Evidence for syllable structure, stress and juncture from segmental durations," J. Phonetics **22**, 283–315.

Bell, A., and Hooper, J. B. (**1978**). "Issues and evidence in syllabic phonology," in *Segments and Syllables*, edited by A. Bell and J. B. Hooper (North–Holland, New York), pp. 9–22.

Burzio, L. (**1994**). *Principles of English Stress* (Cambridge U. P., Cambridge).

Byrd, D. (**1996**). "Influences on articulatory timing in consonant sequences," J. Phonetics **24**, 209–244.

Byrd, D., and Tan, C. C. (**1996**). "Saying consonant clusters quickly," J. Phonetics **24**, 263–282.

Davis, B. L., and MacNeilage, P. F. (**1995**). "The articulatory basis of babbling," J. Speech Hear. Res. **38**, 1199–1211.

Delattre, P. C., Liberman, A. M., and Cooper, F. S. (**1995**). "Acoustic loci and transitional cues for consonants," J. Acoust. Soc. Am. **27**, 769–773.

Delgutte, B., and Kiang, N. Y. S. (**1984**). "Speech coding in the auditory nerve: IV. Sounds with consonantlike dynamic characteristics," J. Acoust. Soc. Am. **75**, 897–907.

Diehl, R. L., and Kluender, K. (**1989**). "On the objects of speech perception," Ecological Psychol. **1**, 121–144.

Diehl, R. L., and Lindblom, B. (**1996**). "Explaining the structure of feature and phoneme inventories," in *Speech Processing in the Auditory System*, edited by S. Greenberg, W. A. Aisnworth, A. Popper, and R. Fay (Springer-Verlag, New York).

Fruchter, D., and Sussman, H. (**1997**). "Perceptual significance of locus equations," J. Acoust. Soc. Am. **105**, 2997–3008.

Greenberg, S. (**1998**). "A syllable-centric framework for the evolution of spoken language," Brain Behav. Sci. **21**, 518.

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

Helfer, K. S., and Huntley, R. A. (**1991**). "Aging and consonant errors in reverberation and noise," J. Acoust. Soc. Am. **90**, 1786–1795.

Henderson, J. B., and Repp, B. (**1982**). "Is a stop consonant released when followed by another stop consonant?" Phonetica **39**, 71–82.

Hoard, J. E. (**1966**). "Juncture and syllable structure in English," Phonetica **15**, 96–109.

Hock, H. H. (**1986**). *Principles of Historical Linguistics* (Mouton de Greater, The Hague).

Kahn, D. (**1976**). "Syllable-based generalizations in English phonology," Ph.D. thesis, MIT.

Krull, D. (**1988**). "Acoustic properties as predictors of perceptual responses: A study of Swedish voiced stops," Phonetic Experimental Research at the Institute of Linguistics, University of Stockholm, PERILUS VII, pp. 66–70.

Liberman, A. M., Delattre, P. C., Gerstman, L. J., and Cooper, F. S. (**1956**). "Tempo of frequency change as a cue for distinguishing classes of speech sounds," J. Exp. Psychol. **52**, 127–137.

Lindblom, B. (**1963**). "On vowel reduction," Report No. 29, The Royal Institute of Technology, Speech Transmission Laboratory, Stockholm, Sweden.

Lindblom, B. (**1986**). "Phonetic universals in vowel systems," in *Experimental Phonology*, edited by J. Ohala and J. Jaeger (Academic, New York), pp. 13–44.

Lindblom, B., MacNeilage, P. F., and Studdert-Kennedy, M. (**1984**). "Self-organizing processes and the explanation of phonological universals," in *Explanations for Language Universals*, edited by B. Butterworth, B. Comrie, and O. Dahl (Mouton, Berlin), pp. 181–203.

Maddieson, I. M. (**1984**). *Patterns of Sounds* (Cambridge U. P., London).

Malecot, A. (**1968**). "The force of articulation of American stops and fricatives as a function of position," Phonetica **18**, 95–102.

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

Moon, S., and Lindblom, B. (**1994**). "Interaction between duration, context, and speaking style in English stressed vowels," J. Acoust. Soc. Am. **96**, 40–55.

Ostreicher, H. J., and Sharf, D. J. (**1976**). "Effects of coarticulation on the identification of deleted consonant and vowel sounds," J. Phonetics **4**, 285–301.

Pickett, J. M., Bunnell, H. T., and Revoile, S. G. (**1995**). "Phonetics of intervocalic consonant perception: retrospect and prospect," Phonetica **52**, 1–40.

Raphael, L. J., and Dorman, M. F. (**1980**). "Silence as a cue to the perception of syllable-initial and syllable-final stop consonants," J. Phonetics **8**, 269–275.

Samuel, A. G. (**1989**). "Insights from a failure of selective adaptation: Syllable-initial and syllable-final consonants are different," Percept. Psychophys. **45**, 485–493.

Sussman, H. M., Bessell, N., Dalston, E., and Majors, T. (**1997**). "An investigation of stop place articulation as a function of syllable position: a locus equation perspective," J. Acoust. Soc. Am. **101**, 2826–2838.

Sussman, H. M., Fruchter, D., Hilbert, J., and Sirosh, J. (**1998**). "Linear correlates in the speech signal: The orderly output constraint," Behav. Brain Sci. **21**, 241–299.

Sussman, H. M., and Shore, J. (**1996**). "Locus equations as phonetic descriptors of consonantal place of articulation," Percept. Psychophys. **58**, 936–946.

Van Wieringen, A. (**1995**). "Perceiving dynamic speechlike sounds: psychoacoustics and speech perception," Ph.D. thesis, Universiteit van Amersterdam, Holland.

Wang, M. D., and Bilger, R. C. (**1973**). "Consonant confusions in noise: a study of perceptual features," J. Acoust. Soc. Am. **54**, 1248–1266.

Willerman, R. (**1994**). "The phonetics of pronouns: articulatory bases of markedness," Ph.D. thesis, University of Texas, Austin.

Winitz, H., Scheib, M. E., and Reeds, J. A. (**1972**). "Identification of stops and vowels for the burst portion of /p,t,k/ isolated from conversational speech," J. Acoust. Soc. Am. **51**, 1309–1317.

Wolf, C. (**1978**). "Perceptual invariance for stop consonants in different positions," Percept. Psychophys. **24**, 315–326.