THE RELATIVE ROLES OF SYNTAX AND PROSODY IN THE PERCEPTION OF THE /š/–/č/ DISTINCTION*

P. J. PRICE

Massachusetts Institute of Technology

and

A. G. LEVITT

Haskins Laboratories and Wellesley College

A silent interval that cues the /š/–/č/ distinction in many contexts is less likely to do so when it coincides with certain boundaries. In natural speech these boundaries are generally marked by both prosody and syntax. We independently varied syntax and prosody to assess their contributions to the phonetic interpretation of silences occurring at these boundaries. We used a set of four sentences, four durations of silence, and two prosodic patterns (Experiment I). We constructed sentences using three techniques that differed in the amount of prosodic control and in naturalness: synthesis by rule, concatenation of naturally produced syllables, and cross-splicing of naturally produced utterances. Silence duration had a strong effect on the perception of the /š/–/č/ contrast in all conditions. For the Synthetic Condition, we also found a strong effect of the prosodic pattern. We found no evidence of any purely syntactic effect. In Experiment II, the two syllables surrounding the silence were excised from the sentences of Experiment I and presented to listeners for labeling. Prosody had a significant effect in the Synthetic Condition and in the Natural Condition. The results indicate that the local prosodic pattern (one syllable with a pitch fall and a greater length) can be sufficient to influence listeners' perception of the /š/–/č/ contrast. There is also evidence that the prosodic information may be subject to context effects.

INTRODUCTION

The introduction of a short silent interval before an appropriate intervocalic fricative noise can change listeners' labelings from "sh" to "ch" (Dorman, Raphael and Liberman, 1979). For example, in the utterance *say shop*, the introduction of silence after the word *say* can change the percept of 'shop' to 'chop.' Rakerd, Dechovitz and Verbrugge (1982) have shown, however, that this change is much less likely to occur when the silence coincides with a sentence boundary. Presumably, the listeners interpret the silence as a consequence of the sentence boundary, i.e., as a pause, rather than as the silence associated with oral closure for /č/. Dechovitz (1980, 1981) has argued that

sentence-internal clause boundaries have a similar effect on listeners' perception and that such boundaries will have an effect even when they are not marked by appropriate prosody.

Syntactic boundaries in natural speech are, however, generally associated with significant prosodic changes that may be largely, or entirely, responsible for the subject's interpretation of the silence. It is therefore important to control carefully for the role of prosody insofar as possible before attributing the effect purely to syntax. Aspects of prosody that may mark clause boundaries include a drop in fundamental frequency (F_0), a lengthening of the clause-final syllable, and a period of silence before the beginning of the next clause. By independently varying the syntax and these prosodic markers in several sentences, we can test the relative roles of syntax and prosody in influencing a listener's decision that the silence is to be attributed to oral closure for /č/ or to a pause followed by /š/.

The separation of syntactic and prosodic effects leads to an important methodological consideration: Since prosody and syntax are often correlated in natural speech, the more effectively the two are separated, the less natural the sentences begin to sound. In our attempt to deal with this problem, we have used three techniques to create the sentences so that some are more natural sounding, but less carefully controlled, and others the reverse.

EXPERIMENT I

Method

Stimuli. Table 1 shows the two pairs of sentences used. Each pair contains an equal number of syllables and shares a large number of words: Sentences 1a and 1b differ, or "are disambiguated," before *pay*, whereas sentences 2a and 2b are disambiguated after *pay*. The two members of each pair differ in syntactic structure: Sentences 1a and 2a have a syntactic break after *pay;* sentences 1b and 2b do not. We used four durations of silence (0, 30, 60, 90 msec) following *pay*. The sentences were generated in two versions: one with a prosodic pattern appropriate for a break following *pay*, and one with a pattern appropriate for no break following *pay*. Patterns appropriate for a break principally involve the syllables immediately before that break. These syllables may show longer duration, a fall or fall-rise F_0 pattern, a tapering off in amplitude, and a following pause. The same syllables occurring in a sentence without such a break are shorter and have flatter F_0 and amplitude patterns. Here we have investigated the combined roles of F_0 pattern and duration in marking the boundary; the two were not separated in this study.

We found in pilot studies that an intervocalic /š/ preceded by silence generally was perceived as /š/ unless the onset was edited to be more abrupt (cf. Rakerd *et al.*, 1982). In order to allow silence to operate as an effective /š/–/č/ cue, we had to edit the friction noise to make it more ambiguous between /š/ and /č/. We shortened the initial friction noise and gave it a sharper rise time. These changes were based on measurements

TABLE 1

Source sentences. Subjects hear either "Shipley" or "Chipley"
following *pay*

1a. Since we have all our back pay, Shipley and I want to leave town.

1b. He wants enough to repay Shipley, and I want to leave town.

2a. That he could pay, Shipley reiterated.

2b. That he could pay Shipley was a shock to me.

of natural speech productions of /č/.[1]

The dependent variable in our design was the perceptual change of 'Shipley' to 'Chipley.' We chose proper names to minimize effects of lexical frequency and semantic expectation. The stressed open syllable *pay* can show clearly the F_0, amplitude and duration patterns that may mark clause finality versus non-finality, and its final high front glide transitions are similar in productions of either *pay ship* or *pay chip*.

The three methods used to create the sentences were: (1) Synthesis by rule: These sentences were not very natural sounding but prosodic patterns were strictly controlled. (2) Concatenation of syllables excised from naturally produced strings: These sentences were more natural than in the Synthetic Condition but prosodic patterns were disrupted. (3) Cross-splicing of large pieces of naturally produced utterances: These sentences sounded natural but prosodic patterns were not strictly controlled.

Synthetic Condition. A version of each of the four sentences in Table 1 was generated using Ingemann's (1978) rules on the OVE-IIIc synthesizer at Haskins Laboratories (Liljencrants, 1968). To facilitate the perceptual change to /č/, the /š/ frication from sentence 1a was edited so that the initial fricative noise was shorter and had a sharper rise time. This frication was used in all the synthetic sentences. Though an intonation "fall" generally occurs in sentence-final position and a "fall-rise" pattern in phrase-final position, the rise part of the fall-rise may occur either before the break or on the first syllable after the break (see, e.g., Cooper and Sorensen, 1981). Delattre (1965) observed that the rise part of the fall-rise pattern is generally not as important in American speech as the fall part. To sort out the relative perceptual values of these two patterns we used two "final" versions of *pay* (one with a fall-rise F_0 pattern and

[1] It is common experimental practice to neutralize cues other than those under investigation. It is somewhat difficult to determine an appropriate neutral value for the /š/−/č/ friction noise. Our pilot studies indicated that what is neutral with respect to /š/−/č/ in utterance-initial position is not neutral in vocalic contexts.
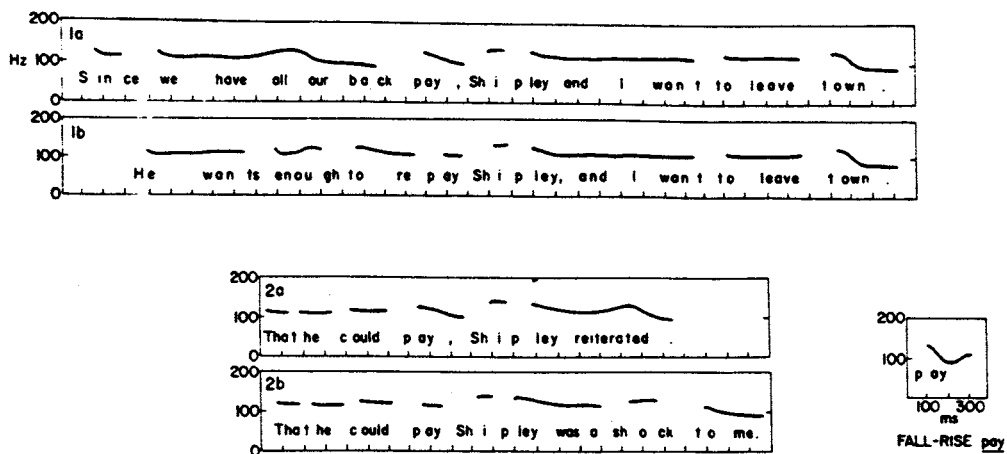
SYNTHETIC CONDITION

Fig. 1. Synthetic Condition sentences with F_0 patterns. The axes at the left show frequency in Hertz. Sentences 1a and 2a contain the *pay* with a fall (final) contour. The flat (non-final) *pay* shown in sentences 2a and 2b was switched with the fall (final) *pay* shown in sentences 1a and 2a in order to control syntax and prosody independently. In the fall-rise part of the Synthetic Condition the F_0 pattern on *pay* shown at the right was substituted everywhere for the fall pattern shown in sentences 1a and 2a at the left. Silence was inserted after *pay*.

the other with a falling F_0 pattern, both of equal length and amplitude), and one "non-final" version (with a shorter duration and a flat amplitude and F_0 pattern). Figure 1 shows the F_0 and temporal patterns for the source sentences used in this condition.

Note that sentence 1b has a syntactic and prosodic break after *Shipley*, while sentence 1a does not. Since this break occurs in the part of the sentence that the two members of the pair are supposed to share, sentence 1a was edited to create a compromise version in which the duration of the final vowel in *Shipley* was increased by 75 msec and an amplitude contour (a symmetric fall and rise) was added. The matching parts of sentences 2a and 2b before *pay* were identical as generated and no editing was necessary.

The Synthetic Condition was divided into two blocks, each consisting of five separate randomizations of 32 stimuli: 4 sentences (see Table 1) × 2 *pays* (final and non-final) × 4 silence durations (0, 30, 60, 90 msec). The two blocks differed in that the final *pay* used had either the fall-rise contour or the fall contour. Digitized versions of each sentence were created before randomization.

Concatenated Condition. In this condition the starting point was natural speech. In order to preserve as much segmental naturalness as possible, while at the same time eliminating most prosodic cues, strings of two or three syllables were recorded, digitized,

TABLE 2

Example of the strings generated for sentence 2b. The middle column contains the syllables that were spliced back together to form the sentence. Sets of strings similar to these 11 strings were generated for the 4 test sentences and for an additional 28 filler sentences. The strings were randomized and read with a list intonation. The pieces in the middle column were then isolated and spliced together to form the sentences in the Concatenated Condition. The symbol ∄ indicates a pause.

∄	that	hat
heet	he	key
key	could	pould
peed	pay	shay
shay	ship	lip
leap	lee	we
we	wuh	zuh
zuh	zuh	shuh
shuh	shock	tock
tuke	to	moo
moo	me	∄

edited and then spliced together to form the sentences in Table 1. A randomized list of these strings was read with list intonation by one of the authors (PJP). The list contained the syllables of the test sentences as well as similar strings from some additional sentences. By list intonation we mean that, in general, all syllables had a pitch fall; the last syllable in a string (the prepausal syllable) fell to a lower level and was longer. An example of the strings used for sentence 2b appears in Table 2.

The middle column in Table 2 contains the syllables to be concatenated with others to form the sentences. Note that the syllable strings were constructed so that the syllables in the middle column were uttered in phonetic contexts similar to that of the part of the sentence into which the syllable was to be spliced. Adjustments of the transcriptions were made to condition phonological rules such as flapping. The syllable strings were low-pass filtered at 5 kHz and sampled at a rate of 10 kHz before editing. One *ship* was used in all the sentences. The friction noise at the beginning was made more ambiguous between /š/ and /č/: it was shortened and its onset made sharper. A single *pay*, from the
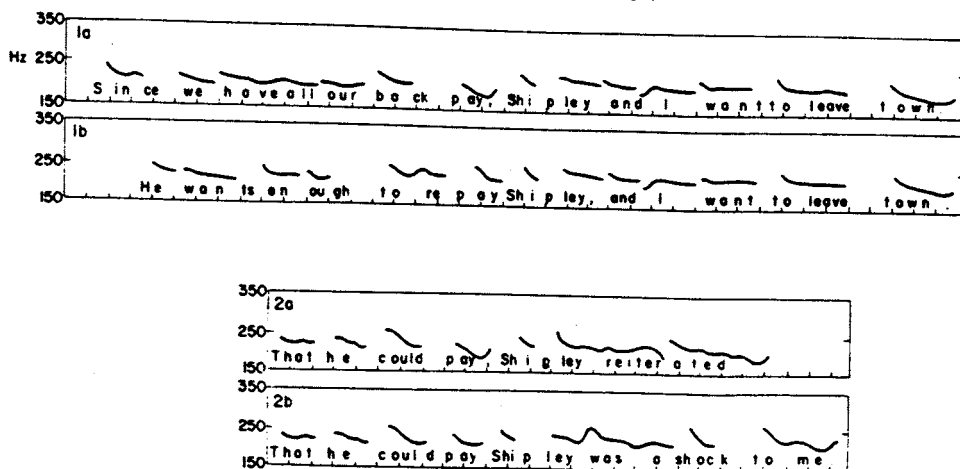
CONCATENATED CONDITION

Fig. 2. Concatenated Condition sentences with F_0 patterns. The axes at the left show frequency in Hertz. The "flattened" (non-final) *pay* shown in sentences 1b and 2b was switched with the original fall-rise (final) *pay* shown in sentences 1a and 2a. Silence was inserted after *pay*.

pre-pausal context, was used. Linear predictive coding (LPC) analysis and resynthesis were used to flatten the F_0 of this syllable, and the waveform editor was used to shorten it from about 300 msec to about 200 msec, thereby creating the "nonfinal" version of *pay*. Analysis of the LPC-flattened *pay* revealed that the F_0 was not flattened during the first 40 to 50 msec of the vowel. This left a sharp F_0 fall at the vocalic onset, which we felt was not unreasonable for a vowel following a voiceless consonant (Hombert, Ohala and Ewan, 1979). The sentences composed of concatenated syllables were edited further to eliminate any audible discontinuities. Figure 2 shows the F_0 and temporal patterns of the source sentences used in this condition.

The four source sentences in Figure 2 were generated in two versions: one with the prepausal *pay* (shown in sentences 1a and 2a) and one with the "flattened" and shortened *pay* (shown in sentences 1b and 2b). We used four durations of silence (0, 30, 60 and 90 msec) between the *pay* and the *ship* of the resulting eight sentences to create 32 stimuli. Five separate randomizations of the 32 stimuli were recorded.

Natural Condition. The four sentences of Table 1 were included in a randomized list containing 28 filler sentences. This list was read by one of the authors (PJP). Sentences with the same words (and, presumably, syntactic structure) but with (presumably) inappropriate prosodic structure were created by cross-splicing pieces of the sentences as indicated in Figure 3. The naturally occurring *ship* in each of the sentences was replaced by the single edited *ship* used in the Concatenated Condition. The resulting eight
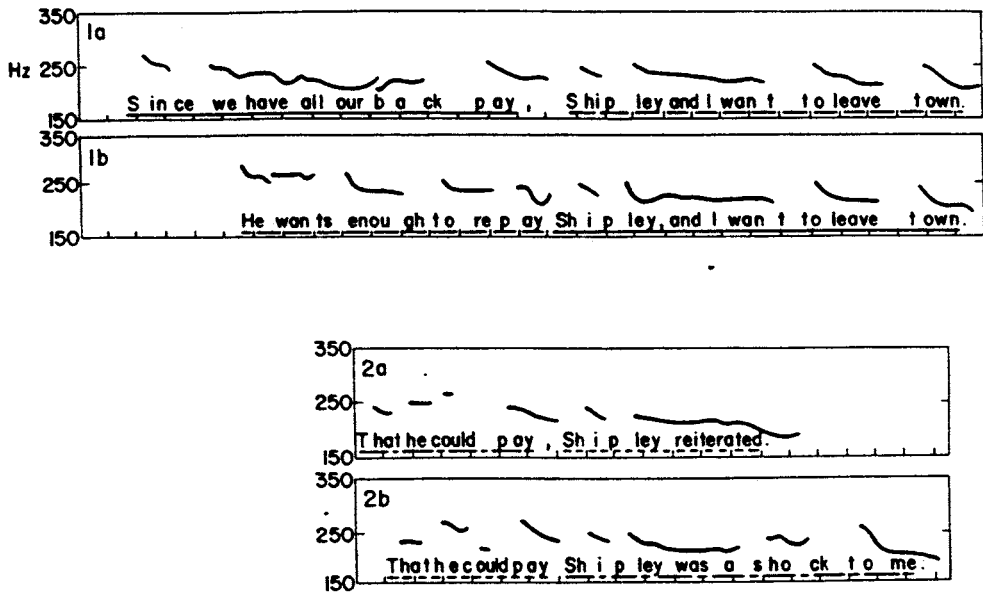
NATURAL CONDITION

Fig. 3. Natural Condition sentences with F_0 patterns. The axes at the left show frequency in Hertz. These sentences were cross-spliced as indicated: Portions of the sentences with the same underlining were joined to form the new sentences. Silence was inserted after *pay*.

sentences were used to generate the 32 stimuli of the experiment (with 0, 30, 60, or 90 msec of silence between *pay* and *Shipley* in each). Again, five separate randomizations of the 32 stimuli were recorded. The F_0 and temporal patterns of the source sentences used in this condition are shown in Figure 3.

Subjects and procedure. Ten Yale undergraduates with no reported history of speech or hearing problems were paid to listen to the four resulting tapes (Synthetic Fall-rise, Synthetic Fall, Concatenated, and Natural) in counter-balanced order over Grason-Stadler model TDH 39-300Z headphones connected to an Ampex tape recorder. Subjects were asked to write "s" if they heard "Shipley" in the sentence and "c" if they heard "Chipley." They were told that it was important to listen to the entire sentence before deciding.

Results

We analyzed the results of the three conditions separately to see whether prosodic pattern, syntactic structure or silence duration affected the number of "s" responses. An analysis of variance was performed on the "s" responses in each of the three conditions

(Synthetic, Concatenated, and Natural). Each analysis included the factors disambiguation (before/after), syntactic context (break/no break), prosody ("final"/"non-final") and silence duration (0, 30, 60 or 90 msec). The analysis of the Synthetic Condition had as an additional factor the pitch change that marked the break after the *pay* (fall/fall-rise).

Figure 4 (thick lines) presents the results of this experiment. The data are averaged across disambiguation (before/after), syntactic context (break/no break), and for the Synthetic Condition, the pitch marker of the break ("fall"/"fall-rise").

In all conditions there was a highly significant main effect of silence duration. In the Synthetic Condition there was also a significant main effect of prosody (final/non-final) (more "s" responses for "final" *pays*, as expected): $F(1, 9) = 12.34, p = 0.0066$, and a significant interaction of prosody and silence duration, $F(3, 27) = 12.84, p < 0.0001$. There was no significant difference between the two final pitch patterns on *pay* ("fall" versus "fall-rise"). Finally, there was a significant interaction of syntax (break/no break) and prosody (final/non-final), $F(1, 9) = 6.45, p = 0.0318$. When there was a "final" contour on *pay*, sentences with a syntactic break received slightly more "s" responses than sentences without such a break; whereas when there was a "non-final" contour on *pay*, sentences with a syntactic break had slightly fewer "s" responses than sentences without a break. The expectation, of course, would be that if syntax were an independent cue to the listener, sentences with a "final" contour on *pay* and a syntactic break would show more "s" responses than would sentences without a syntactic break. This is the opposite of what was obtained. A significant interaction, $F(3, 27) = 3.51, p = 0.0287$, between F_0 change (fall, fall-rise) and silence duration was due to the fact that the number of "s" responses was sometimes slightly higher for fall, other times higher for fall-rise, depending on the given silence duration. There was no systematic pattern of differences. A significant three-way interaction of syntax, prosody and disambiguation, $F(1, 9) = 5.95, p = 0.0374$ was also present: for a "final" contour on *pay* there were slightly more "s" responses for sentence 1a (with a break) than for sentence 1b (without a break), whereas there were slightly more "s" responses for sentence 2b (without a break) than for sentence 2a (with a break).

In the Concatenated Condition the only significant effect, besides that of silence duration ($F(3, 27) = 39.66, p < 0.0001$), was a three-way interaction among disambiguation (before/after), prosody (final/non-final), and silence duration, $F(3, 27) = 3.08$, $p = 0.0445$, due to the fact that sentences disambiguated before the break showed a slight rise in number of "s" responses for the "final" *pays* at the longest silence duration whereas sentences disambiguated after the break did not. Although there was no significant prosodic effect, seven of the ten subjects did show more "s" responses when the preceding *pay* had the "final" prosodic pattern as opposed to the "non-final" one.

In the Natural Condition the effect of silence duration was most pronounced: subjects' responses changed almost completely from "s" to "c" with the introduction of 30 msec of silence, regardless of sentence type. There was also a significant interaction of disambiguation (before/after) and prosody (final/non-final), $F(1, 9) = 9.00, p = 0.0150$: for the sentences disambiguated in their initial part, subjects showed a greater number of "s" responses for a "final" *pay* than for a "non-final" *pay*. However, in the sentence
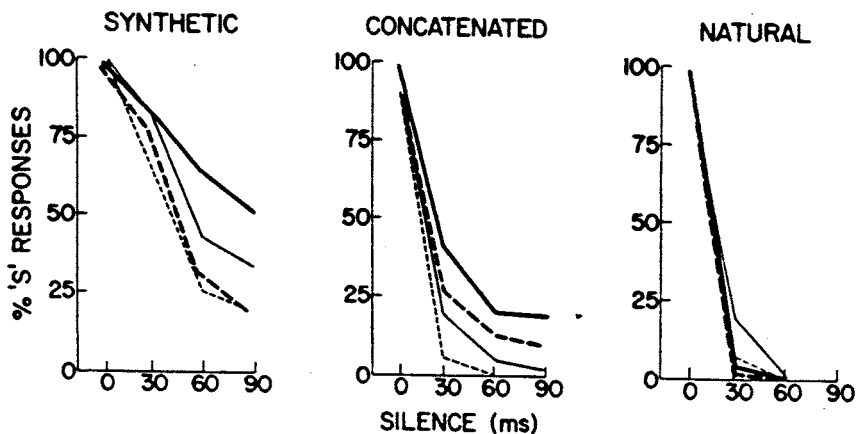
Fig. 4.   Percent "s" responses are plotted for the Synthetic (left), Concatenated (middle) and Natural (right) Conditions. Thick lines represent responses to sentences; thin lines to controls. Solid lines show responses to items with "final" *pay*s; dashed lines to "non-final" *pay*s.

pair that was disambiguated in its final part, subjects showed a greater number of "s" responses for a non-final *pay*. There was another significant two-way interaction ($F(3, 27)$ = 3.86, p = 0.0203) of prosody (final/non-final) and silence duration, and a significant three-way interaction ($F(3, 27)$ = 5.77, p = 0.0035) of those factors and disambiguation (before/after) which both seem due to the fact that each of the four individual *pay*s used in this condition produced slightly different cross-over points. Experiment II deals with this issue more directly.

Discussion

There was a clear phonetic effect of silence duration in all three conditions. For all subjects and all conditions the introduction of silence after *pay* caused subjects to report "Chipley" rather than "Shipley."

The effect of the F_0 and duration patterns of *pay* (final/non-final) on the number of "s" ("Shipley") responses is clear in the case of the synthetically produced sentences. Subjects' responses to the two final F_0 patterns on *pay* (fall and fall-rise), however, did not differ, which suggests that both were equally good at signalling a break to American English listeners. In the Natural Condition a prosodic effect was obtained only in the sentences that were disambiguated before the syntactic break but not in the sentences disambiguated after the syntactic break. However, since the sentences in the Synthetic Condition that were disambiguated after the break show a prosodic effect, we believe that the failure to find one in the sentences disambiguated after the break in the Natural Condition reflects the fact that the overall prosodic contours of the natural sentences

are not as well controlled as in the other conditions. We do not believe that site of disambiguation was the crucial factor. Finally, the Concatenated Condition showed a trend in the direction of an effect of prosodic pattern.

We found no evidence of a purely syntactic effect: Grammatical structure of the sentences independent of the prosodic patterns was not a significant factor, nor was there a trend in that direction for any of our three conditions. A negative result, of course, does not prove that such syntactic effects cannot occur. However, it is essential to disentangle possible prosodic effects from syntactic effects in order to clearly demonstrate the latter. Our results show that prosody can play an important role in the perception of the /ʒ/−/č/ distinction.

What, then, is the domain of the prosodic effect? Is the falling pitch pattern and longer duration of the *pay* sufficient to cue a change in the number of "s" responses regardless of context? Experiment II addresses these questions.

EXPERIMENT II

Method

The various *pay ship*s from the preceding experiment were isolated by waveform editing. For the Synthetic Condition, this resulted in three *pay*s (two "final" − fall and fall-rise, and one "non-final," which was flat in F_0 and shortened) times four durations of silence, or 12 source stimuli. For the Concatenated Condition, the two *pay*s (the original pre-pausal and its flattened and shortened version) and four silence durations resulted in eight stimuli. For the Natural Condition the four *pay*s (one for each of the sentences in Figure 1) and four silence durations resulted in 16 stimuli. Ten randomizations of each set of stimuli were prepared, blocked by condition, and presented for labeling to 12 new subjects in counterbalanced order. Subjects were asked to write "s" if they heard "pay ship" and "c" if they heard "pay chip."

Results

A two-way analysis of variance (prosody and silence duration) was performed on each of the conditions. In all three conditions, silence duration was highly significant. Prosody was a significant main effect in the Synthetic Condition (as in Experiment I), $F(2, 22) = 5.23$, $p = 0.0138$, and in the Natural Condition, $F(1, 11) = 6.34$, $p = 0.0286$ (unlike Experiment I where it was part of a significant interaction), but not in the Concatenated Condition. There was an interaction of prosody and silence duration in the Synthetic Condition, $F(6, 66) = 2.25$, $p = 0.0489$, and in the Natural Condition, $F(3, 33) = 4.71$, $p = 0.0076$.

Figure 4 (thin lines) shows the results of Experiment II. As before, results for the Synthetic Condition are averaged over the two final versions of *pay* used (fall/fall-rise), and results for the Natural Condition are averaged over the two tokens of the final *pay*s and over the two tokens of the non-final *pay*s.

In order to compare Experiments I and II, we did an unequal N analysis of variance

on the results of the two experiments for each condition.

For the Synthetic Condition, the combined analysis (Experiments I and II) showed highly significant effects of prosody ($F(2, 40) = 15.82, p < 0.0001$) and silence duration ($F(3, 60) = 67.02, p < 0.001$) and a highly significant interaction of prosody and silence duration ($F(6, 120) = 7.51, p < 0.0001$) as had been found in each of the separate analyses. There was also a significant three-way interaction of task, prosody and silence duration, $F(6, 120) = 4.07, p = 0.0009$, reflecting a greater number of "s" responses for silence durations of 60 msec or greater in the sentences than in the *pay ships*.

For the Concatenated Condition, we found a highly significant effect of silence duration ($F(3, 60) = 187.05, p < 0.0001$), the only significant effect in each of the separate analyses, and a significant interaction of task and silence duration ($F(3, 60) = 3.49, p = 0.0211$), again showing a greater number of "s" responses for the longer silence durations (here, 30 msec or longer) in the sentences than in the *pay ships*.

For the Natural Condition, we found in the combined analysis a significant effect of silence duration ($F(3, 60) = 1035.57, p < 0.0001$), as we had in the separate analyses, and a significant prosodic effect ($F(1, 20) = 6.16, p = 0.0221$) as well as a significant interation of prosody and silence duration ($F(3, 60) = 6.14, p = 0.001$), as we had in Experiment II.

Discussion

In the separate analysis of the results of Experiment II alone, a strong effect of silence duration was again demonstrated in each of the three conditions. In the Synthetic Condition, as in the previous experiment, prosody was a significant main effect and a significant interactive effect with silence duration. In the Concatenated Condition prosody was not a significant effect in the sentences or in the controls. The original *pay* in this condition was from a prepausal context. Although the F_0 was flattened by LPC analysis and resynthesis and the syllable was shortened, other cues to "finality" may have remained. It is also possible that the syllable was insufficiently flattened in F_0 and/or shortened. In any case, though the flattening and shortening resulted in something more like a non-final *pay,* as seen by the trends in the data, the effect did not reach significance. In the Natural Condition, prosody as a main effect and its interaction with silence duration were both significant in the controls, though they were not in the sentences (Experiment I).

When we compare the results of Experiments I and II, task emerges as a significant interactive effect in both the Synthetic and the Concatenated Conditions. In both cases the interactions appear due to the fact that in the experiment with sentences, there tend to be a greater number of "s" responses at longer silence durations than in the experiment with the two syllables. In the Synthetic and the Concatenated Conditions prosody is more controlled, but the sentences sound less natural and less coarticulated. It seems reasonable that subjects might interpret silence as a random pause (and not as closure for the affricate) in these less natural sounding sentences and therefore respond with more "s" responses. The lack of naturalness would be less salient in the experiment with the two syllables. Furthermore, since the utterances are shorter and are less likely to be heard as sentences, the silences may be less likely to be interpreted as pauses.

In sum, we see a very similar pattern of results for the two experiments. A sharp F_0 fall and a longer duration seem to be sufficient to sway listener judgments towards pause plus /š/ rather than /č/, when other factors are neutralized.[2] Further, the more effectively these factors are neutralized, i.e. in the Synthetic Condition, the more important these aspects of prosody can be. Of course, in actual speech communication such factors are not generally separated. That people can make reliable judgments when they are is evidence, we feel, that they are a significant part of the many factors people are attuned to in speech understanding.

GENERAL DISCUSSION·

The results of these experiments show a clear pattern of the effect of prosody on the perception of the /š/–/č/ distinction in a variety of contexts. Although no purely syntactic effects were found here, it is possible that a change in the subject's task would elicit such an effect. Miller (1982), for example, has suggested that variations in prosody (speaking rate, in her case) are "automatically" taken into account by the listener, whereas semantic effects only emerge when the task focuses on meaning. Semantic or syntactic structures are more likely to play a role when the task more directly demands them. We also believe that, in general, listeners use any strategies and any information available (see also Cutler, 1982). We would argue, however, that prosody is more available to the listener as an aid in initial parsing of a sentence than syntax can be at this stage.

Our data also provide evidence for the importance of the syllables immediately adjacent to the boundary in cueing that boundary. The same *ship* was used in the Concatenated and Natural Conditions, yet the patterns of "s" responses differ. Some context effects of domains larger than this are suggested in the comparisons of the two experiments.

A further result of our study bears on methodology. We feel that the cross-splicing of large pieces of naturally produced sentences is the least appropriate of the techniques we used. The fact that in these sentences the key parameters are sometimes conflicting and in general not independently controlled make the data difficult to interpret. On the other hand, naturalness is a highly desirable feature in test stimuli.

There is much evidence that the pitch contour and temporal properties of the local environment of a break can carry a great deal of weight in marking that break (see. e.g., Cooper and Sorensen, 1977; Larkey, 1980; Pierrehumbert, 1980; Grosjean, 1982). We found that these factors can outweigh those of the syntax and semantics of a sentence. This, together with other reports of segmental and suprasegmental interactions (see, e.g., Klatt and Cooper, 1975; Lehiste, 1975; Summerfield, 1975; Nooteboom and Doodeman, 1980), suggest the possibility that listeners may use suprasegmental information to assign an initial syntactic structure before decoding the rest of the information. We see research along these lines as promising for investigations of acoustic correlates of prosodic

---

2   That listeners continued to hear /š/ even when the edited friction noise was preceded by short intervals of silence indicates that we did not in editing eliminate all the cues that identify /š/.

information and of their role in marking perceptual units for the listener.

REFERENCES

COOPER, W.E. and SORENSEN, J.M. (1977). Fundamental frequency contours at syntactic boundaries. *Journal of the Acoustical Society of America*, **62**, 683-692.

COOPER, W.E. and SORENSEN, J.M. (1981). *Fundamental Frequency in Sentence Production* (New York).

CUTLER, A. (1982). Prosody and sentence perception in English. In J. Mehler, E. Walker and M. Garrett (eds.), *Perspectives on Mental Representation* (Hillsdale, N.J.), pp. 201-260.

DECHOVITZ, D.R. (1980). Constraints imposed on phonetic processing by syntax. *Journal of the Acoustical Society of America*, **67**, S99 (Abstract).

DECHOVITZ, D.R. (1981). Phonetic integration as a potential test for controversial surface structures. *Journal of the Acoustical Society of America*, **69**, S115 (Abstract).

DELATTRE, P. (1965). *Comparing the Phonetic Features of English, French, German, and Spanish: An Interim Report* (Philadelphia).

DORMAN, M.F., RAPHAEL, L.J. and LIBERMAN, A.M. (1979). Some experiments on the sound of silence in phonetic perception. *Journal of the Acoustical Society of America*, **65**, 1518-1532.

GROSJEAN, F. (1982). How long is the sentence? Prediction and prosody in the on-line processing of language. Unpublished ms.

HOMBERT, J.-M., OHALA, J.J. and EWAN, W.G. (1979). Phonetic explanations for the development of tones. *Language*, **55**, 37-58.

INGEMANN, F. (1978). Speech synthesis by rule using the FOVE program, *Haskins Laboratories Status Report on Speech Research*, SR-54, 165-174.

KLATT, D.H. and COOPER, W.E. (1975). Perception of segment duration in sentence contexts. In A. Cohen and S. Nooteboom (eds.), *Structure and Process in Speech Perception* (New York), pp. 68-89.

LARKEY, L.S. (1980). The role of prosodic information in speech perception (University Microfilms).

LEHISTE, I. (1975). The phonetic structure of paragraphs. In A. Cohen and S. Nooteboom (eds.), *Structure and Process in Speech Perception* (New York), pp. 195-206.

LILJENCRANTS, J. (1968). The OVEIII speech synthesizer, *I.E.E.E. Transactions: Audio*, **16**, 137-140.

MILLER, J. (1982). On the distinction between prosodic and semantic factors in word identification. *Journal of the Acoustical Society of America*, **71**, S95 (Abstract).

NOOTEBOOM, S.G. and DOODEMAN, G.J.N. (1980). Production and perception of vowel length in spoken sentences. *Journal of the Acoustical Society of America*, **67**, 276-287.

PIERREHUMBERT, J. (1980). The phonology and phonetics of English intonation. Ph.D. dissertation, Massachusetts Institute of Technology.

RAKERD, B., DECHOVITZ, D.R. and VERBRUGGE, R. (1982). An effect of sentence finality on the phonetic significance of silence. *Language and Speech*, **25**, 267-282.

SUMMERFIELD, A.Q. (1975). How a full account of segmental perception depends on prosody and vice versa. In A. Cohen and S. Nooteboom (eds.), *Structure and Process in Speech Perception* (New York), pp. 51-68.

ADDRESSES OF CONTRIBUTORS

DENES, Peter B., American Telephone and Telegraph, Room 2EO, 225 Littleton Road, Morris Plains, NJ 07950, U.S.A.

GELFER, Dr. Carole E., Haskins Laboratories, 270 Crown Street, New Haven, CT 06511, U.S.A.

HIEKE, Dr. Adolf E., Sprach-und Literaturwissenschaftliche Fakultät, Universität Bayreuth, Postfach 3008, 8580 Bayreuth, West Germany.

KOWAL, Dr. Sabine, Siemensstrasse 13, 1000 Berlin 46, West Germany.

LEVITT, Dr. A.G., Haskins Laboratories, 270 Crown Street, New Haven, CT 06511, U.S.A.

MASTERSON, Dr. J.F., School of Remedial Linguistics, University of Dublin, Trinity College, Dublin, Ireland.

MCGARR, Dr. Nancy S., Haskins Laboratories, 270 Crown Street, New Haven, CT 06511, U.S.A.

MCMANUS, Dr. I.C., Department of Psychology, Bedford College, University of London, Regent's Park, London NW1 4NS, U.K.

MULLINS, Mr. E., Department of Statistics, University of Dublin, Trinity College, Dublin 2, Ireland.

MULVIHILL, Ms. A.K., School of Remedial Linguistics, University of Dublin, Trinity College, Dublin 2, Ireland.

NELSON, Professor Marie Wilson, Department of English, George Mason University, Fairfax, VA 22030, U.S.A.

O'CONNELL, Professor Daniel C., S.J., Loyola University of Chicago, 6525 No. Sheridan Road, Chicago, IL 60626, U.S.A.

POWER, Dr. M.J., Munro Clinic, Guy's Hospital, St. Thomas Street, London S.E.1, U.K.

PRICE, Dr. P.J., Research Laboratory of Electronics, 36-511, Massachusetts Institute of Technology, Cambridge, MA 02139, U.S.A.

PRINZ, Dr. Philip M., Division of Education and Communication Disorders, The Pennsylvania State University, 217 Moore Building, University Park, PA 16802, U.S.A.

RUBIN, Professor Donald L., Department of Speech Communication and Language Education, The University of Georgia, Athens, GA 30602, U.S.A.

PUBLICATIONS RECEIVED

Bennis, Hans and Van Lessen Kloeke (eds.), *Linguistics in The Netherlands 1983*, Publications in Language Science 12, (Foris Publications, Dordrecht-Holland/Cinnaminson-U.S.A., 1983).

Dik, Simon C. (ed.), *Advances in Functional Grammar*, Publications in Language Sciences 11, (Foris Publications, Dordrecht-Holland/Cinnaminson-U.S.A., 1983).

Dimmendaal, Jan, *The Turkana Language*, Publications in African Languages and Linguistics, (Foris Publications, Dordrecht-Holland/Cinnaminson-U.S.A., 1983).

Journal of Experimental Psychology: Human Perception and Performance, 9 (1983), 4.

Journal of Experimental Psychology: Learning, Memory, and Cognition, 9 (1983), 3.

Leuvense Bijdragen, Tijdschrift voor Germaanse Filologie, 72 (1983), 2.

Van der Hulst, Harry and Smith, Norval (eds.), *The Structure of Phonological Representations (Part II)*, Linguistic Models 3, (Foris Publications, Dordrecht-Holland/Cinnaminson-U.S.A., 1982).