# Phoneme-monitoring in the context of different phonetic sequences

**Anne Cutler\* and William £. Cooper**

*Massachusetts Institute of Technology, Cambridge Massachusetts* 02139, *U.S.A.*

*Received 26th July* 1977

**Abstract:** The order of some conjoined words is rigidly fixed (e.g. *dribs and drabs/\*drabs and dribs).* Both phonetic and semantic factors can play a role in determining the fixed order. An experiment was conducted to test whether listeners' reaction times for monitoring a predetermined phoneme are influenced by phonetic constraints on ordering. Two such constraints were investigated: monosyllable-bisyllable and high-low vowel sequences. In English, conjoined words occur in such sequences with much greater frequency than their converses, other factors being equal. Reaction times were significantly shorter for phoneme monitoring in monosyllable-bisyllable sequences than in bisyllable-monosyllable sequences. However, reaction times were not significantly different for high-low vs. low-high vowel sequences.

## Introduction

When two words are conjoined, their order of appearance is typically free. Thus, both *book and lamp* and *lamp and book* are grammatical. In other cases, however, the order of conjoined words is fixed. Examples include *this and that/\*that and this, kith and kin/ \*kin and kith, spic and span/\*span and spic,* and *free and easy/\*easy and free.* Both semantic and phonetic factors determine the ordering (Abraham, 1950; Cooper & Ross, 1975; Jespersen, 1961; Malkiel, 1959). For English, words that appear in the first position of fixed-order conjoinings tend to refer to semantic concepts such as Here, Now, Adult, Male, Positive, and Singular, whereas words appearing in second position tend to refer to opposite concepts. In terms of sound structure, words in first position tend to be monosyllables and contain a word-initial segment that is low in obstruency, as well as word-medial high vowel (Cooper and Ross, 1975). Many of the semantic and phonetic rules for determining the order of conjoined words also apply to hyphenated word combinations, such as *razzle-dazzle, flim-flam,* and *super-duper.*

This study was aimed at the question of whether some of the phonetic constraints on ordering serve to aid listeners in processing sequential material consisting of nonsense syllables. Two phonetic factors were studied: monosyllables-bisyllables, and high-low vowels.

The tendency for conjuncts in first position to be monosyllabic is considered to be the strongest of the phonetic determinants of ordering; in cases where conjoinings pit two independently motivated phonetic factors against one another, the monosyllabic constraint typically overrides, as in *boots and saddles, bread and butter,* and *rough and ready*

(Cooper & Ross, 1975). The tendency for conjuncts in first position to contain a high vowel is also observed in English with regularity, particularly in the case of the high vowel /I/, as in *wig-wag, flim-flam, pitter-patter, shilly-shally, tick-tock,* and *sing-song.*

A phoneme-monitoring task was used to test whether listeners' reaction times to a phoneme target would be faster when nonsense sequences obeyed the ordering tendencies noted than when they did not. This task was chosen because it appears to be responsive to on-line processing difficulty during speech perception; reaction times are sensitive to the occurrence of a low-frequency word or an ambiguous word immediately prior to the target-bearing word, as well as to the preceding intonation contour (Foss, 1969, 1970; Cutler, 1976; Ref. Note 1). The ease or difficulty of perceptual processing at the phonetic level presumably contributes to on-line processing difficulty and should therefore affect reaction time in the phoneme-monitoring task. It was thus predicted that, if the phonetic constraints on fixed-order conjoinings found in English serve to aid perceptual processing, listeners' reaction times to phoneme targets should be faster when the preceding material obeys these constraints.

## Methods

### Subjects

Twenty-four members of the M.I.T. community, recruited by advertisements on campus, participated in the experiment. They were each paid $2 for their participation.

### Materials

(i) *Vowel height.* Sixteen lists of nonsense syllables, of either 8 or 10 syllables in length were constructed. In eight of these, the vowel sequence was /I/-/ae/-/I/-/ae/; in the other eight the sequence was reversed. An example list is: mab jiv zam fip bav vid kag nis shabhin (phoneme target /k/). In half of the lists of each sequence, the target-bearing item was a syllable with the vowel /I/; in the other half it was a syllable with the vowel /ae/. These sixteen lists, along with twenty-eight filler lists (consisting of syllables with vowels other than /ae/ or /I/), were recorded by a male speaker of standard American. The intonation of the lists was kept as monotonous as possible. By splicing off the first syllable of each experimental list, a further sixteen lists were created, whereby each previously /ae/-/I/ sequence became an /I/-/ae/ sequence, and vice versa.[1] Thus the final set contained thirty-two experimental lists plus the twenty-eight filler lists.

Measurements of syllable duration were made of a random sample of the lists; the mean duration of /ae/ syllables in the lists sampled was 510 ms, the mean duration of /I/ syllables 462 ms, and the mean inter-syllable interval 559 ms.

(ii) *Syllable structure.* Thirty-two lists were constructed, sixteen of which consisted of monosyllable-bisyllable sequences (dib-pibble, zit-gittle), the other sixteen of bisyllable-monosyllable sequences (pibble-dib, gittle-zit). The bisyllables were stressed on the first syllable and carried somewhat greater stress than the monosyllables in each case. In half of each group the target-bearing item was a monosyllable, in the other half a

[1]The spliced lists were thus each one syllable shorter than their unspliced counterparts. However, the target-bearing syllable never occurred within the first 4 syllables in an experimental list, even after splicing. The syllables removed were not spliced on to the end of the lists to preserve length, since the target-bearing syllable was also never the last or 2nd to last syllable of a list, and the RTs were such that the subject had always pressed the button by the time the end of the list was reached.

bisyllable. The lists were recorded in random order by the same speaker as for the other lists; the thirty-two experimental lists were preceded by four lists of similar structure.

Measurements of a random sample of lists in this condition established a mean duration for bisyllables in the sample of 344 ms, a mean duration for monosyllables of 255 ms, a mean interval between two items of a pair of 93 ms and a mean interval between pairs of 297 ms.

### *Procedure*

Subjects were tested in groups of up to three at a time. The lists were presented binaurally over Telephonics TDH-49 headphones, and the subjects were instructed to press a response button as soon as they heard a syllable beginning with the target for that list. The target sound, which could be either /k/ or /d/, was specified for each list immediately prior to the beginning of the list.

A signal (inaudible to the subjects) on the tape, synchronized with the onset of the target sound, started a separate digital timer for each subject, which was stopped by the subject pressing the button. The RTs were recorded by the experimenter.

Each set of lists was presented separately, and the subjects were given a rest period after each set. All subjects heard the Vowel Height lists before the Syllable Structure lists.

## Results

The mean reaction time for each subject was computed for each of the experimental conditions. The average reaction times across all twenty-four subjects are presented in Table I.

Table. I  Mean reaction time scores (ms)

| | Vowel height | | |
| | /ae/ | /I/ | $\bar{X}$ |
| --- | --- | --- | --- |
| /I/-/ae/ sequence | 412 | 405 | 408 |
| /ae/-/I/ sequence | 388 | 419 | 404 |
| $\bar{X}$ | 400 | 412 | |

| | Syllable structure | | |
| | Monosyllable | Bisyllable | $\bar{X}$ |
| --- | --- | --- | --- |
| Monosyllable–bisyllable | 488 | 495 | 492 |
| Bisyllable–monosyllable | 572 | 508 | 540 |
| $\bar{X}$ | 530 | 502 | |

The results from each set of lists were subjected to 2 separate analyses of variance, in one of which Subjects was treated as a random factor, while in the other the random factor was Items.

### *Vowel height*

Neither main effect (Sequence; Vowel Height of Target-Bearing Item) was significant in either analysis; however, the interaction between these two effects was significant in the analysis by Subjects ($F_{(1,23)}$ = 5.30, $P < 0.05$), with /I/ syllables producing faster reaction times than /ae/ syllables in the /I/-/ae/ sequences, but /ae/ syllables producing faster reaction times than /I/ syllables in the /ae/-/I/ sequences. Since this result did not hold up in the analysis by Items ($P > 0.2$), it is doubtful whether any importance should be attached to it.

### *Syllable structure*

The main effect for Sequence was significant in both analyses (Subjects: $F_{(1,23)}$ = 7.33, $P < 0.02$; items: $F_{(1,7)}$ = 5.62, $P < 0.05$), with monosyllable-bisyllable sequences eliciting faster reaction times than bisyllable-monosyllable sequences. Neither monosyllables nor bisyllables produced reaction times which were significantly faster on either analysis, and the interaction was again non-significant in both analyses.

**Discussion**

The results of the Syllable Structure condition support the hypothesis that phonetic information in monosyllable-bisyllable sequences is easier to process than the same information contained in bisyllable-monosyllable sequences. On the other hand, the null result obtained in the Vowel Height condition provided no support for the hypothesis that the vowel sequencing constraint serves as an aid to perceptual processing. However, it must be noted that the syllables in this condition were presented at equal intervals rather than as pairs analogous to the hyphenated expressions found in language; the failure to obtain the predicted difference may perhaps be attributable to this fact. In addition, the insertion of filler items within the vowel test may have prevented listeners from being consciously aware of the presence of vowel sequencing, whereas sequencing was readily apparent in the other condition.

A subsidiary result, tangential to the main issue of this study, should also be mentioned; namely, the failure to find a reaction time difference between stressed and unstressed items in the Syllable Structure condition, in which bisyllables, which bore heavier stress, did not elicit faster reaction times than monosyllables. This result is in accord with the findings of Shields, McHugh & Martin (1974) that stressed nonsense syllables produced faster reaction times than unstressed when they were embedded in a sentence context but not when they were presented in a list of other nonsense syllables.

In conclusion, then, evidence has been provided that a phonetic constraint on syllable structure, observed in conjoined items in English, can facilitate perceptual processing of speech. The variety of such constraints, and the exact manner in which they function in the perceptual process, provide topics for further research.

## References

Abraham, R. M. (1950). Fixed order of coordinates: A study in lexicography. *Modern Language Journal* 34, 276-287.

Cooper, W. E. & Ross, J. R. (1975). World order. In *Papers from the Parasession on Functionalism* (Grossman, R. E., San, L. J. & Vance, T. J., eds.) pp. 63-111. Chicago: Chicago Linguistic Society.

Cutler, A. (1976). Phoneme-monitoring reaction time as a function of preceding intonation contour. *Perception & Psychophysics* **20,** 55-60.

Foss, D. J. (1969). Decision processes during sentence comprehension: Effects of lexical item difficulty and position upon decision times. *Journal of Verbal Learning and Verbal Behavior* 8, 457-462.

Foss, D. J. (1970). Some effects of ambiguity upon sentence comprehension. *Journal of Verbal Learning and Verbal Behavior* 9, 699-706.

Jespersen, O. (1961). *A Modern English Grammar on Historical Principles. Part VI Morphology.* London: George Allen and Unwin Ltd.

Malkiel, Y. (1959). Studies in irreversible binomials. *Lingua* 8, 113-160.

Shields, J. L., McHugh, A. & Martin, J. G. (1974). Reaction time to phoneme targets as a function of rhythmic cues in continuous speech. *Journal of Experimental Psychology* **102,** 250-255.

*Note*

Cutler, A. The nature of a reaction time task: phoneme-monitoring for example. Paper presented at the colloquium on reaction time methodology in sentence processing experiments, International Congress of Psychology, Paris, 1976.