

---

# INTELLIGIBILITY OF SYLLABLE-TIED INTERRUPTED SPEECH

A. V. KATWIJK—J. T. HART\*

Without intending to contribute to the issue of the theoretical status of the syllable, we would like to study the perceptual implications of Pilch's<sup>1</sup> taking the syllable as a frame within which distributional constraints can be described. Furthermore, we were struck by Huggins'<sup>2</sup> observation, that intelligibility of disrupted speech was minimal when interruptions occurred at a rate that he believed to interfere with the syllabic structure. In our experiments, using syllable-tied interruptions, to ensure that of each syllable a comparable fragment will be audible, we tried to find out which is the part of the syllable that is most vulnerable to mutilation and what is the assumed effect of distortion of the rhythmic structure of syllabic sequences. In a first experiment, described elsewhere,<sup>3</sup> synchronization between syllables and interruptions was brought about by having the text spoken in a metrical way. Although locations of audible and suppressed fragments were not known exactly, the results strongly suggested, that the onset of the syllable, i.e. (C<sub>1</sub>) + V, contributes more to perception than the coda, i.e. V + (C<sub>2</sub>). As regards distributional restrictions, these very onsets constitute the distributionally free, or less predictable parts of syllables. This makes the scores, viz. 84 and 89 % in situations with audible onsets and 58 and 25 % with onsets suppressed, understandable.

Also, transgression of the distributional syllable boundary seemed to make the task more difficult for subjects.

If we now hypothesize the syllable to be a unit of intelligibility, and if we further assume that the perceptual syllable boundary coincides with the distributional one, we may expect that any addition, in the presentation of fragments, of parts of adjacent syllables, would make it less easy to allow one and only one syllable identification per presented fragment, whereas any such addition would also distort the rhythmic pattern, thus accumulating the difficulties.

The experiment to be described here is meant to test the hypotheses mentioned,

---

\* IPO, Eindhoven, Holland.

<sup>1</sup> H. Pilch, *Phonemtheorie I*, Basel 1964.

<sup>2</sup> A. W. F. Huggins, Distortion of the Temporal Pattern of Speech, *JASA*, 36, 1055—1064.

<sup>3</sup> A. van Katwijk, On Perceptual Units in Speech, *IPO Annual Progress Report 1*, 1966.

while avoiding an exaggeration of the rhythmic structure in the metrically spoken text, and securing a higher accuracy in locations of interruptions.

In preparation, a continuous text of 1100 syllables was read aloud by the experimenter at a controlled speech rate, and recorded. With a segmentator, C—V links were established and in re-recording, 120 ms before these links, impulses were put on the second track. In the experiment, these impulses triggered a time delay device, which controlled the operation of the gate with which the interruption was brought about, in six different situations, always with fixed durations of 110 ms (i.e. half the average syllable length) either in the audible, or in the suppressed parts. The six situations were (in their relation to the C—V link): 1. —35 to +75; 2. —95 to +15; 3. +30 to +140 ms audible, and 4. +15 to +125; 5. —80 to +30; 6. —35 to +75 ms suppressed. 18 Subjects were instructed to try to reconstruct the entire text, using the technique of shadowing. Training took place with 88 syllables, interrupted every 220 ms with an on-off ratio of unity. Subjects were presented with all six situations in a cyclically changing order.

The results, in percentage correct of the total number of syllables (with standard error of the mean), are: 1.  $70 \pm 2$ , 2.  $65 \pm 3$ , 3.  $46 \pm 3$ , 4.  $42 \pm 4$ , 5.  $34 \pm 2$ , and 6.  $15 \pm 2\%$ .

This outcome in the first place amply confirms the main effect in the first experiment, viz. that the onsets give the greatest contribution to intelligibility. But this

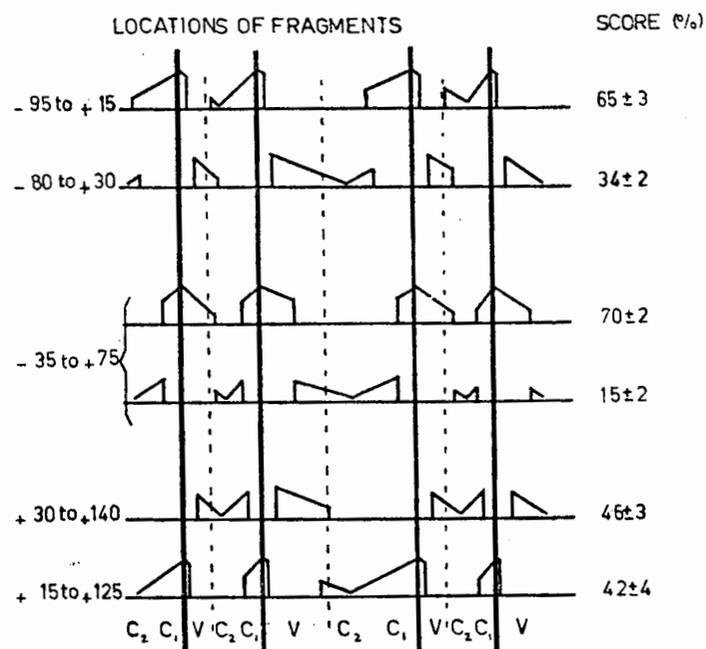


Fig. 1. Illustration of the six interruption situations used in the second experiment, for an imaginary fragment of the text. Left: opening and closing moments in relation to the C-V links: middle: audible fragments; right: mean scores and standard errors of the means.

cannot explain the difference in scores of situations 1 and 2 on the one hand and 4 on the other. We should, therefore, consider the two hypothesized conditions together, viz. a) audibility of the C—V link and b) non-interference with the rhythmic structure. It appears then, that in situations 1 and 2 both conditions are met, whereas situation 6 does not fulfil either of them. Situation 4, which allows for audibility of the C—V link, suffers from severe distortion of the rhythm, as both beginning and end of the audible part have no fixed time relation to the C—V link of the syllable from which it originates. Situation 3, which is the reverse of 4 in that it meets condition b, but not a, has about the same score as 4, showing that the two conditions approximately compensate each other. Situation 5, compared with 3, meets condition b to a limited extent in having only the beginning of the audible part related to the appropriate C—V link.

Conclusion: We feel justified in accepting the syllable as a unit of intelligibility with respect to 1. the onset, whose contribution is correlated with its distributional properties, and 2. the rhythmic pattern in syllabic sequences, whose contribution, though indirect, is well specifiable in terms of time relations with the onset.

## DISCUSSION

### *Pulgram:*

On what criteria did experimenters cut the utterance into syllables, that is, how did they know where to make the cut-off point? Is not recognizability of truncated unit due to redundancy of language rather than to distribution of phonemes? If nonsense syllables with proper Dutch phoneme distribution had been used, would not the identification possibilities have been reduced?

### *'t Hart:*

ad Pulgram: There was no need to cut the utterance into syllables; we only established C-V links.

It was our very aim to examine *where* redundancy is situated. That redundancy is not homogeneously distributed, and that there is some correlation with phoneme distribution is one of the results. Extra redundancy in our material raises the overall score, but cannot be held responsible for the differences between the various interruption situations.